(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau

(43) International Publication Date
23 September 2004 (23.09.2004)

PCT

(10) International Publication Number
WO 2004/080162 A2

(51) International Patent Classification[7]: **A01K**

(21) International Application Number:
PCT/US2004/006378

(22) International Filing Date: 1 March 2004 (01.03.2004)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/453,126	7 March 2003 (07.03.2003)	US
60/490,452	28 July 2003 (28.07.2003)	US
60/536,677	15 January 2004 (15.01.2004)	US

(71) Applicant *(for all designated States except US)*: **AVI-GENICS, INC.** [US/US]; 111 Riverbend Road, Georgia Biobusiness Center, Athens, GA 30605 (US).

(72) Inventors: **HARVEY, Alex**; 188 Sunset Drive, Athens, GA 30606 (US). **LEAVITT, Markley, C.**; 1140 Latham Drive, Watkinsville, GA 30677 (US). **CHRISTMANN,**

Leandro; 1311 Victoria Road, Watkinsville, GA 30677 (US).

(74) Agent: **HAYZER, David, J.**; Womble Carlyle Sandridge & Rice, PLLC, P.O. Box 7037, Atlanta, GA 30357-0037 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK,

(54) Title: INTEGRASE MEDIATED AVIAN TRANSGENESIS
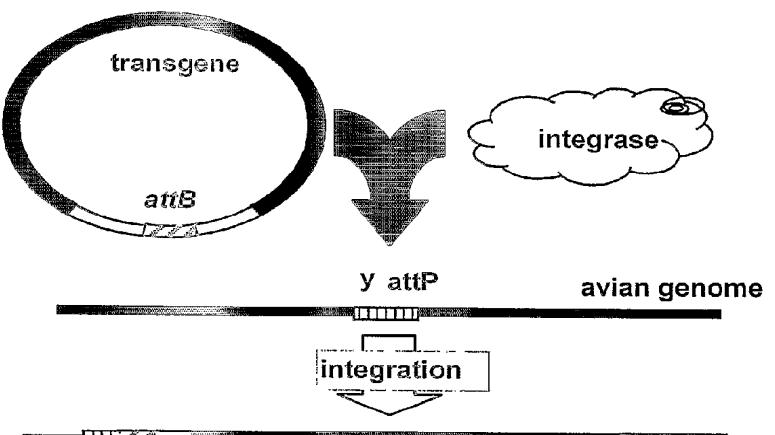
(57) Abstract: The invention provides methods for integrating a heterologous polynucleotide into the genome of an avian cell. The methods deliver to an avian cell a polynucleotide and a source of integrase activity that mediates recombination between the polynucleotides and the genomic DNA of the avian cell. The invention provides modified avian or artificial chromosomes as vectors to shuttle transgenes or gene clusters into an avian genome. Another aspect of the invention are avian cells genetically modified with a transgene vector. One cell line for the delivery and integration of a transgene comprises a heterologous attP site and, optionally, a region for expressing the integrase. Methods are also included for the production of a heterologous polypeptide by transgenic avian tissue involve integrating a heterologous polynucleotide into the avian genome. Subsequently, a mature transgenic avian is derived by transferring the transgenic blastodermal cells to an embryo and allowing that embryo to develop fully.

TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Title of the Invention

INTEGRASE MEDIATED AVIAN TRANSGENESIS

5   The present application claims priority from U.S. provisional patent application Serial Nos. 60/453,126, filed March 7, 2003, 60/490,452, filed July 28, 2003, and 60/536,677 filed January 15, 2004 and which are hereby incorporated by reference in their entireties.

10   ### Field of the Invention

The present invention relates to the field of biotechnology, and more specifically to the field of avian genome modification. Disclosed herein are compositions, vectors, and methods of use thereof, for the generation of genetically transformed avian cells and transgenic birds.

15

Background

Transgenic technology to convert animals into "bioreactors" for the production of specific proteins or other substances of pharmaceutical interest (Gordon *et al.,* 1987, *Biotechnology* 5: 1183-1187; Wilmut *et al.,* 1990, *Theriogenology* 33: 113-123)
20   offers significant advantages over more conventional methods of protein production by gene expression. Recombinant nucleic acid molecules, for instance, have been engineered and incorporated into transgenic animals so that an expressed heterologous protein may be joined to a protein or peptide that allows secretion of the transgenic expression product into milk or urine, from which the protein may then be recovered.
25   These procedures, however, may require lactating animals, with the attendant costs of maintaining individual animals or herds of large species, such as cows, sheep, or goats.

Historically, transgenic animals have been produced almost exclusively by microinjection of the fertilized egg. The pronuclei of fertilized eggs are microinjected
30   *in vitro* with foreign, i.e., xenogeneic or allogeneic, heterologous DNA or hybrid

DNA molecules. The microinjected fertilized eggs are then transferred to the genital tract of a pseudopregnant female (e.g., Krimpenfort *et al.,* U.S. Pat. No. 5,175,384).

One system that holds potential is the avian reproductive system. The production of an avian egg begins with formation of a large yolk in the ovary of the

5   hen. The unfertilized oocyte or ovum is positioned on top of the yolk sac. After ovulation, the ovum passes into the infundibulum of the oviduct where it is fertilized if sperm are present, and then moves into the magnum of the oviduct, which is lined with tubular gland cells. These cells secrete the egg-white proteins, including ovalbumin, lysozyme, ovomucoid, conalbumin and ovomucin, into the lumen of the

10   magnum where they are deposited onto the avian embryo and yolk. The hen oviduct offers outstanding potential as a protein bioreactor because of the high levels of protein production, the promise of proper folding and post-translation modification of the target protein, the ease of product recovery, and the shorter developmental period of chickens compared to other potential animal species.

15   One method for creating permanent genomic modification of an eukaryotic cell is to integrate an introduced DNA into an existing chromosome. Only retroviruses have so far provided efficient integration. However, retroviral integration is directed to a number, albeit limited, of insertion sites within the recipient genome so that positional variation in heterologous gene expression can be evident.

20   Unpredictability as to which insertion site is targeted introduces an undesirable lack of control over the procedure. An additional limitation of the use of retroviruses is that the size of the nucleic acid molecule encoding the virus and heterologous sequences is restricted to about 8 kb. Although wild-type adeno-associated virus (AAV) often integrates at a specific region in the human genome, vectors derived from AAV do not

25   integrate site-specifically due to the deletion of the toxic *rep* gene. Other well-known methods for genomic modification of animal cells include transfection of DNA using calcium phosphate co-precipitation, electroporation, lipofection, microinjection, protoplast fusion and particle bombardment, all of which methods typically produce random integration and at low frequency. Homologous recombination produces site-

30   specific integration, but the frequency of such integration usually is very low.

An alternative method that has been considered for driving the integration of heterologous nucleic acid fragments into a chromosome is the use of a site-specific recombinase (integrase) that can catalyze the insertion or excision of nucleic acid fragments. These enzymes recognize relatively short unique nucleic acid sequences

5   that serve for both recognition and recombination. Examples include Cre (Sternberg & Hamilton, 1981, *J. Mol. Biol.* 150: 467-486, 1981), Flp (Broach *et al.*, 1982, *Cell* 29: 227-234, 1982) and R (Matsuzaki *et al.*, 1990, *J. Bact.* 172: 610-618, 1990).

A novel class of phage integrases that includes the integrase from the phage phiC31 can mediate highly efficient integration of transgenes in mammalian cells both

10   *in vitro* and *in vivo* (Thyagarajan *et al.*, *Mol. Cell Biol.* 21: 3926-3934 (2001)). Constructs and methods of using recombinase to integrate heterologous DNA into a plant, insect or mammalian genome are described by *Calos* in U.S. Patent Serial No. 6,632,672.

The phiC31 integrase is a member of a subclass of integrases, termed serine

15   recombinases, that include R4 and TP901-1. Unlike the phage lambda integrases, which belong to a tyrosine class of recombinases, the serine integrases do not require cofactors such as integration host factor. The phiC31 integrase normally mediates integration of the phiC31 bacteriophage into the genome of *Streptomyces* via recombination between the attP recognition sequence of the phage genome and the

20   attB recognition sequence within the bacterial genome. When a plasmid is equipped with a single attB site, phiC31 integrase will detect and mediate crossover between the attB site and a pseudo-attP site within the mammalian genome. Such pseudo-attP integration sites have now been identified in the mouse and human genomes. If the heterologous DNA is in a circular or supercoiled form, the entire plasmid becomes

25   integrated with *attL* and *attR* arms flanking the nucleic acid insert. PhiC31 integrase is not able to mediate the integration into genomic DNA of sequences bearing attP sites.

PhiC31 integrase-mediated integration results in the destruction of the recognition or recombination sites themselves so that the integration reaction is

30   irreversible. This will bypass the primary concern inherent with other recombinases, i.e., the reversibility of the integration reaction and excision of the inserted DNA.

It has been estimated that there are 50 to 100 pseudo-attP sites in mammalian genomes (mouse and human) and some sites are apparently preferred for integration over others. The chicken genome, however, is only about one-third the size of mammalian genomes, and it was unknown whether there would be a sufficient

5   number of pseudo attP sites in the chicken genome to allow efficient integrase-mediated integration.

We have found that the phiC31 integrase is active in avian cells, increasing the rate of integration over that of a non-integrase-mediated integration. Furthermore, we have determined that the phiC31 integrase works well at both 37° Celsius and 41°

10   Celsius, showing that it will function in the environment of a developing avian embryo.

A need still exists, however, for methods by which avian chromosomes can be permanently modified in an efficient and site-specific manner and the genetically transformed cells used to generate transgenic birds.

15

Summary of the Invention

Integration of a transgene into a defined chromosomal site is useful to improve the predictability of expression of the transgene, which is particularly advantageous when creating transgenic avians. Transgenesis by methods that randomly insert a

20   transgene into an avian genome is often inefficient since the transgene may not be expressed at the desired levels or in desired tissues.

A novel class of phage integrases, and in particular the integrase from phage phiC31, can mediate the efficient integration of transgenes into target cells both *in vitro* and *in vivo*. When a plasmid is equipped with a single attB site, phiC31

25   integrase detects attP homologous sequences, termed pseudo-attP sites, in a target genome and mediates crossover between the attB site and a pseudo attP site.

The present invention provides novel methods and recombinant polynucleotide molecules for transfecting and integrating a heterologous nucleic acid molecule into the genome of an avian cell. The methods of the invention deliver to an avian cell

30   population a first nucleic acid molecule that comprises a region encoding a bacterial

recombination site. A source of integrase activity also delivered top the avian cell can be an integrase-encoding nucleic acid sequence and its associated promoter included in the first nucleic acid molecule or as a region of a second nucleic acid molecule that may be co-delivered with the polynucleotide molecule. Alternatively, integrase protein itself can be delivered directly to the target cell.

The recombinant nucleic acid molecules of the present invention may further comprise a heterologous nucleotide sequence operably linked to a promoter so that the heterologous nucleotide sequence, when integrated into the genome DNA of a recipient avian cell, can be expressed to yield a desired polypeptide. The nucleic acid molecule may also include a second transcription initiation site, such as an internal ribosome entry site (IRES), operably linked to a second heterologous polypeptide-encoding region desired to be expressed with the first polypeptide in the same cell.

The heterologous nucleic acid molecule of the present invention may include a cassette for the expression in a recipient avian cell of a desired heterologous polypeptide. Optionally, the nucleic acid molecules may further comprise a marker such as, but not limited to, a puromycin resistance gene, a luciferase gene, EGFP-encoding gene, and the like.

Once delivered to a recipient avian cell, the phiC31 integrase mediates recombination between the att site within the nucleic acid molecule and a bacteriophage attachment site within the genomic DNA of the avian cell. Both att sites are disrupted and the nucleic acid molecule, with partial att sequences at each end, is stably integrated into the genome attP site. The phiC31 integrase, by disrupting the att sites of the incoming nucleic acid and of the recipient site within the avian cell genome, precludes any subsequent reverse recombination event that would excise the integrated nucleic acid and reduce the overall efficiency of stable incorporation of the heterologous nucleic acid.

Following delivery of the nucleic acid molecule and a source of integrase activity into an avian cell population and integrase-mediated recombination, the cells may be returned to an embryo. Late stage blastodermal cells may be returned to a hard shell egg, which is resealed for incubation until hatching. Stage I cells may be directly microinjected with the polynucleotide and source of integrase activity, or

isolated, transfected and returned to a stage I embryo which is reimplanted into a hen for further development. Alternatively, the transfected cells may be maintained in *in vitro* culture.

The present invention further provides modified isolated avian or artificial chromosomes useful as vectors to shuttle transgenes or gene clusters into the avian genome. By delivery to the modified chromosome to an isolated recipient cell, the target cell, and progeny thereof, become trisomic. The additional or trisomic chromosome will not affect the subsequent development of the recipient cell and/or an embryo, nor interfere with the reproductive capacity of an adult bird developed from such cells or embryos. The chromosome will also be stable within chicken cells. The invention provides methods to isolate a population of chromosomes for delivery into chicken embryos or early cells.

The method comprises inserting a lac-operator sequence into an isolated chromosome and, optionally, inserting a desired transgene sequence within the same chromosome. The lac operator region is typically a concatamer of a plurality of lac operators for the binding of multiple lac repressor molecules. A recombinant DNA molecule is constructed that includes an identified region of the target chromosome, a recombination site such as attB or attP, and the lac-operator concatamer. The recombinant molecule is delivered to an avian cell, and homologous recombination will integrate the heterologous polynucleotide and the lac-operator concatamer into the targeted chromosome. A tag-polypeptide, such as the GPF-lac-repressor fusion protein, binds to the lac-operator sequence for identification and isolation of the genetically modified chromosome. The tagged mitotic chromosome can be isolated using, for instance, flow cytometry.

Another aspect of the present invention is an avian cell genetically modified with a transgene vector by the methods of the invention. For example, in one embodiment, the transformed cell can be a chicken early stage blastodermal cell or a genetically transformed cell line, including a sustainable cell line. The transfected cell may comprise a transgene stably integrated into the nuclear genome of the recipient cell, thereby replicating with the cell so that each progeny cell receives a copy of the transfected nucleic acid. A particularly useful cell line for the delivery and integration

of a transgene comprises a heterologous attP site that can increase the efficiency of integration of a polynucleotide by phiC31 integrase and, optionally, a region for expressing the integrase.

Another aspect of the present invention is methods of expressing a heterologous polypeptide in an avian cell by stably transfecting a cell by using site-specific integrase-mediation and a recombinant nucleic acid molecule, as described above, and culturing the transfected cell under conditions suitable for expression of the heterologous polypeptide under the control of the avian transcriptional regulatory region.

Yet another aspect of the present invention concerns transgenic birds, such as chickens, comprising a recombinant nucleic acid molecule and which preferably (though optionally) express a heterologous gene in one or more cells in the animal. Embodiments of the methods for the production of a heterologous polypeptide by the avian tissue involve providing a suitable vector and introducing the vector into embryonic blastodermal cells together with an integrase, preferably phiC31 integrase, so that the vector can integrate into the avian genome. A subsequent step involves deriving a mature transgenic avian from the transgenic blastodermal cells by transferring the transgenic blastodermal cells to an embryo and allowing that embryo to develop fully, so that the cells become incorporated into the bird as the embryo is allowed to develop. An alternative is to transfer a transfected nucleus to an enucleated recipient cell which may then develop into a zygote and ultimately an adult bird. The resulting chick is then grown to maturity.

In various embodiments of the transgenic bird of the present invention, the expression of the transgene may be restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, *trans*-acting factors acting on the transcriptional regulatory region operably linked to the polypeptide-encoding region of interest of the present invention and which control gene expression in the desired pattern. Tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example,

conditional recombination systems or prokaryotic transcriptional regulatory sequences.

The invention can be used to express, in large yields and at low cost, a wide range of desired proteins including those used as human and animal pharmaceuticals,

5   diagnostics, and livestock feed additives. Proteins such as growth hormones, cytokines, structural proteins and enzymes including human growth hormone, interferon, lysozyme, and β-casein are examples of proteins which are desirably expressed in the oviduct and deposited in eggs according to the invention.

Additional objects and aspects of the present invention will become more

10   apparent upon review of the detailed description set forth below when taken in conjunction with the accompanying figures, which are briefly described as follows.

Brief Description of the Figures

Fig. 1 illustrates phage integrase-mediated integration. A plasmid vector

15   bearing the transgene includes the attB recognition sequence for the phage integrase. The vector along with integrase-coding mRNA, a vector expressing the integrase, or the integrase protein itself, are delivered into cells or embryos. The integrase recognizes DNA sequences in the avian genome similar to attP sites, termed pseudo-attP, and mediates recombination between the attB and pseudo-attP sites, resulting in

20   the permanent integration of the transgene into the avian genome.

Fig. 2 illustrates the persistent expression of luciferase from a nucleic acid molecule after phiC31 integrase-mediated integration into chicken cells.

Fig. 3 illustrates the results of a puromycin resistance assay to measure phiC31 integrase-mediated integration into chicken cells.

25   Fig. 4 illustrates phiC31 integrase-mediated integration into quail cells. Puromycin resistance vectors bearing attB sites were cotransfected with phiC31 integrase, or a control vector, into QT6 cells, a quail fibrosarcoma cell line. One day after transfection, puromycin was added. Puromycin resistant colonies were counted 12 days post-transfection.

30   Figs. 5A and 5B illustrate that phiC31 integrase can facilitate multiple integrations per avian cell. A puromycin resistance vector bearing an attB site was

cotransfected with an enhanced green fluorescent protein (EGFP) expression vector bearing an attB site, and a phiC31 integrase expression vector. After puromycin selection, many puromycin resistant colonies expressed EGFP in all of their cells. Figs. 5A and 5B are the same field of view with EGFP illuminated with ultraviolet

5   light (Fig. 5A) and puromycin resistant colonies photographed in visible light (Fig. 5B). In Fig. 5B, there are 4 puromycin resistant colonies, two of which are juxtaposed at the top. One of these colonies expressed EGFP.

Fig. 6 shows maps of the small vectors used for integrase assays.

Fig. 7 shows integrase promotes efficient integration of large transgenes in

10   avian cells.

Fig. 8 shows maps of large vectors used for integrase assays.

Fig. 9 illustrates the nucleotide sequence of the integrase-expressing plasmid pCMV-31int (SEQ ID NO: 1).

Fig. 10 illustrates the nucleotide sequence of the plasmid pCMV-luc-attB

15   (SEQ ID NO: 2).

Fig. 11 illustrates the nucleotide sequence of the plasmid pCMV-luc-attP (SEQ ID NO: 3).

Fig. 12 illustrates the nucleotide sequence of the plasmid pCMV-pur-attB (SEQ ID NO: 4).

20   Fig. 13 illustrates the nucleotide sequence of the plasmid pCMV-pur-attP (SEQ ID NO: 5).

Fig. 14 illustrates the nucleotide sequence of the plasmid pCMV-EGFP-attB (SEQ ID NO: 6).

Fig. 15 illustrates the nucleotide sequence of the plasmid p12.0-lys-

25   LSPIPNMM-CMV-pur-attB (SEQ ID NO: 7).

Fig. 16 illustrates the nucleotide sequence of the plasmid pOMIFN-Ins-CMV-pur-attB (SEQ ID NO: 8).

Fig. 17 illustrates the nucleotide sequence of the integrase-expressing plasmid pRSV-Int (SEQ ID NO: 9).

30   Fig. 18 illustrates the nucleotide sequence of the plasmid pCR-XL-TOPO-CMV-pur-attB (SEQ ID NO: 10).

Fig. 19 illustrates the nucleotide sequence of the attP containing polynucleotide SEQ ID NO: 11.

Fig. 20 illustrates in schematic form the integration of a heterologous att recombination site into an isolated chromosome. The attB sequence is linked to selectable maker such as a puromycin expression cassette and is flanked by sequences found in the target site of the chromosome to be modified. The DNA is transfected into cells containing the chromosome and stable transfectants are selected by drug resistance. Site specific integration may be confirmed by several techniques including PCR.

Fig. 21 illustrates the persistent expression of luciferase from a nucleic acid molecule after phiC31 integrase-mediated integration into chicken cells bearing a wild-type attP sequence.

Detailed Description of the Preferred Embodiments

This description uses gene nomenclature accepted by the Cucurbit Genetics Cooperative as it appears in the *Cucurbit Genetics Cooperative Report* 18:85 (1995), which are incorporated herein by reference in its entirety. Using this gene nomenclature, genes are symbolized by italicized Roman letters. If a mutant gene is recessive to the normal type, then the symbol and name of the mutant gene appear in italicized lower case letters.

The disclosures of publications, patents, and published patent specifications referenced in this application are hereby incorporated by reference into the present disclosure to more fully describe the state of the art to which this invention pertains.

Definitions

For convenience, definitions of certain terms employed in the specification, examples, and appended claims are collected here.

As used in this specification and the appended claims, the singular forms "a," "an" and "the" include plural references unless the content clearly dictates otherwise. Thus, for example, reference to "an antigen" includes a mixture of two or more such agents.

The term "avian" as used herein refers to any species, subspecies or race of organism of the taxonomic class *ava*, such as, but not limited to chicken, turkey, duck, goose, quail, pheasants, parrots, finches, hawks, crows and ratites including ostrich, emu and cassowary. The term includes the various known strains of *Gallus gallus*, or

5 chickens, (for example, White Leghorn, Brown Leghorn, Barred-Rock, Sussex, New Hampshire, Rhode Island, Australorp, Minorca, Amrox, California Gray), as well as strains of turkeys, pheasants, quails, duck, ostriches and other poultry commonly bred in commercial quantities. It also includes an individual avian organism in all stages of development, including embryonic and fetal stages. The term "avian" also may

10 denote "pertaining to a bird", such as "an avian (bird) cell."

The term "nucleic acid" as used herein refers to any natural or synthetic linear and sequential array of nucleotides and nucleosides, for example cDNA, genomic DNA, mRNA, tRNA, oligonucleotides, oligonucleosides and derivatives thereof. For ease of discussion, such nucleic acids may be collectively referred to herein as

15 "constructs," "plasmids," or "vectors." The term "nucleic acid" further includes modified or derivatized nucleotides and nucleosides such as, but not limited to, halogenated nucleotides such as, but not only, 5-bromouracil, and derivatised nucleotides such as biotin-labeled nucleotides.

The terms "polynucleotide," "oligonucleotide," and "nucleic acid sequence"

20 are used interchangeably herein and include, but are not limited to, coding sequences (polynucleotide(s) or nucleic acid sequence(s) which are transcribed and translated into polypeptide *in vitro* or *in vivo* when placed under the control of appropriate regulatory or control sequences); control sequences (e.g., translational start and stop codons, promoter sequences, ribosome binding sites, polyadenylation signals,

25 transcription factor binding sites, transcription termination sequences, upstream and downstream regulatory domains, enhancers, silencers, and the like); and regulatory sequences (DNA sequences to which a transcription factor(s) binds and alters the activity of a gene's promoter either positively (induction) or negatively (repression)). No limitation as to length or to synthetic origin are suggested by the terms described

30 above.

As used herein the terms "peptide," "polypeptide" and "protein" refer to a

polymer of amino acids in a serial array, linked through peptide bonds. A "peptide" typically is a polymer of at least two to about 30 amino acids linked in a serial array by peptide bonds. The term "polypeptide" includes proteins, protein fragments, protein analogues, oligopeptides and the like. The term "polypeptides" contemplates

5   polypeptides as defined above that are encoded by nucleic acids, produced through recombinant technology (isolated from an appropriate source such as a bird), or synthesized. The term "polypeptides" further contemplates polypeptides as defined above that include chemically modified amino acids or amino acids covalently or noncovalently linked to labeling moieties.

10   The terms "percent sequence identity" or "percent sequence similarity" as used herein refer to the degree of sequence identity between two nucleic acid sequences or two amino acid sequences as determined using the algorithm of Karlin & Attschul, *Proc. Natl. Acad. Sci.* 87: 2264-2268 (1990), modified as in Karlin & Attschul, *Proc. Natl. Acad. Sci.* 90: 5873-5877 (1993). Such an algorithm is incorporated into the

15   NBLAST and XBLAST programs of Attschul *et al.*, 1990, *T. Mol. Biol.* Q15: 403-410. BLAST nucleotide searches are performed with the NBLAST program, score = 100, word length = 12, to obtain nucleotide sequences homologous to a nucleic acid molecule of the invention. BLAST protein searches are performed with the XBLAST program, score = 50, word length = 3, to obtain amino acid sequences homologous to

20   a reference polypeptide. To obtain gapped alignments for comparison purposes, Gapped BLAST is utilized as described in Attschul *et al.*, *Nucl. Acids Res.* 25: 3389-3402 (1997). When utilizing BLAST and Gapped BLAST programs, the default parameters of the respective programs (e.g. XBLAST and NBLAST) are used. Other algorithms, programs and default settings may also be suitable such as, but not only,

25   the GCG-Sequence Analysis Package of the U.K. Human Genome Mapping Project Resource Centre that includes programs for nucleotide or amino acid sequence comparisons. Examples of preferred algorithms are FASTA and BESTFIT.

The terms "recombinant nucleic acid" and "recombinant DNA" as used herein refer to combinations of at least two nucleic acid sequences that are not naturally

30   found in a eukaryotic or prokaryotic cell. The nucleic acid sequences may include, but are not limited to, nucleic acid vectors, gene expression regulatory elements, origins

of replication, suitable gene sequences that when expressed confer antibiotic resistance, protein-encoding sequences and the like. The term "recombinant polypeptide" is meant to include a polypeptide produced by recombinant DNA techniques. A recombinant polypeptide may be distinct from a naturally occurring

5   polypeptide either in its location, purity or structure. Generally, a recombinant polypeptide will be present in a cell in an amount different from that normally observed in nature.

The term "gene" or "genes" as used herein refers to nucleic acid sequences that encode genetic information for the synthesis of a whole RNA, a whole protein, or any

10   portion of such whole RNA or whole protein. Genes that are not naturally part of a particular organism's genome are referred to as "foreign genes," "heterologous genes" or "exogenous genes" and genes that are naturally a part of a particular organism's genome are referred to as "endogenous genes". The term "gene product" refers to an RNA or protein that is encoded by the gene. "Endogenous gene products" are RNAs

15   or proteins encoded by endogenous genes. "Heterologous gene products" are RNAs or proteins encoded by "foreign, heterologous or exogenous genes" and are, therefore, not naturally expressed in the cell.

The term "expressed" or "expression" as used herein refers to the transcription from a gene to give an RNA nucleic acid molecule at least complementary in part to a

20   region of one of the two nucleic acid strands of the gene. The term "expressed" or "expression" as used herein may also refer to the translation from an RNA molecule to give a protein, a polypeptide or a portion thereof.

The term "operably linked" refers to an arrangement of elements wherein the components so described are configured so as to perform their usual function. Control

25   sequences operably linked to a coding sequence are capable of effecting the expression of the coding sequence. The control sequences need not be contiguous with the coding sequence, so long as they function to direct the expression thereof. For example, intervening untranslated yet transcribed sequences can be present between a promoter sequence and the coding sequence and the promoter sequence can still be

30   considered "operably linked" to the coding sequence.

The term "transcription regulatory sequences" as used herein refers to nucleotide sequences that are associated with a gene nucleic acid sequence and which regulate the transcriptional expression of the gene. Exemplary transcription regulatory sequences include enhancer elements, hormone response elements, steroid response elements, negative regulatory elements, and the like.

5

The term "promoter" as used herein refers to the DNA sequence that determines the site of transcription initiation by an RNA polymerase. A "promoter-proximal element" is a regulatory sequence generally within about 200 base pairs of the transcription start site.

10

The term "internal ribosome entry sites (IRES)" as used herein refers to a region of a nucleic acid, most typically an RNA molecule, wherein eukaryotic initiation of protein synthesis occurs far downstream of the 5' end of the RNA molecule. A 43S pre-initiation complex comprising the elf2 protein bound to GTP and Met-tRNA$_i^{Met}$, the 40S ribosomal subunit, and factors elf3 and 3lf1A may bind to

15

an "IRES" before locating an AUG start codon. An "IRES" may be used to initiate translation of a second coding region downstream of a first coding region, wherein each coding region is expressed individually, but under the initial control of a single upstream promoter. An "IRES" may be located in a eukaryotic cellular mRNA.

The term "coding region" as used herein refers to a continuous linear

20

arrangement of nucleotides which may be translated into a polypeptide. A full length coding region is translated into a full length protein; that is, a complete protein as would be translated in its natural state absent any post-translational modifications. A full length coding region may also include any leader protein sequence or any other region of the protein that may be excised naturally from the translated protein.

25

The terms "vector" or "nucleic acid vector" as used herein refer to a natural or synthetic single or double stranded plasmid or viral nucleic acid molecule (RNA or DNA) that can be transfected or transformed into cells and replicate independently of, or within, the host cell genome. The term "expression vector" as used herein refers to a nucleic acid vector that comprises a transcription regulatory region operably linked

30

to a site wherein is, or can be, inserted, a nucleotide sequence to be transcribed and,

optionally, to be expressed, for instance, but not limited to, a sequence coding at least one polypeptide.

The term "transfection" as used herein refers to the process of inserting a nucleic acid into a host cell. Many techniques are well known to those skilled in the

5   art to facilitate transfection of a nucleic acid into an eukaryotic cell. These methods include, for instance, treating the cells with high concentrations of salt such as a calcium or magnesium salt, an electric field, detergent, or liposome mediated transfection, to render the host cell competent for the uptake of the nucleic acid molecules, and by such methods as micro-injection into a pro-nucleus, sperm-

10   mediated and restriction-mediated integration.

The terms "recombinant cell" and "genetically transformed cell" refer to a cell comprising a combination of nucleic acid segments not found in a single cell with each other in nature. A new combination of nucleic acid segments can be introduced into an organism using a wide array of nucleic acid manipulation techniques available

15   to those skilled in the art. The recombinant cell may harbor a vector that is extragenomic, i.e. that does not covalently insert into the cellular genome, including a non-nuclear (e.g. mitochondrial) genome(s). A recombinant cell may further harbor a vector or a portion thereof that is intragenomic, i.e. covalently incorporated within the genome of the recombinant cell.

20   As used herein, a "transgenic avian" is any avian, as defined above, including the chicken and quail, in which one or more of the cells of the avian contain heterologous nucleic acid introduced by manipulation, such as by transgenic techniques. The nucleic acid may be introduced into a cell, directly or indirectly, by introduction into a precursor of the cell by way of deliberate genetic manipulation,

25   such as by microinjection or by infection with a recombinant virus. Genetic manipulation also includes classical cross-breeding, or *in vitro* fertilization. A recombinant DNA molecule may be integrated within a chromosome, or it may be extrachromosomally replicating DNA.

The terms "chimeric animal" or "mosaic animal" are used herein to refer to

30   animals in which the recombinant gene is found, or in which the recombinant is expressed, in some but not all cells of the animal. The term "tissue-specific chimeric

animal" indicates that the recombinant gene is present and/or expressed in some tissues but not others.

As used herein, the term "transgene" means a nucleic acid sequence that is partly or entirely heterologous, i.e., foreign, to the transgenic animal or cell into which

5   it is introduced, or, is homologous to an endogenous gene of the transgenic animal or cell into which it is introduced, but which is designed to be inserted, or is inserted, into the animal's genome in such a way as to alter the genome of the cell into which it is inserted (e.g., it is inserted at a location which differs from that of the natural gene or its insertion results in a knockout).

10   The term "cytokine" as used herein refers to any secreted polypeptide that affects a function of cells and modulates an interaction between cells in the immune, inflammatory or hematopoietic response. A cytokine includes, but is not limited to, monokines and lymphokines. Examples of cytokines include, but are not limited to, interferon α2b, Interleukin-1 (IL-1), Interleukin-6 (IL-6), Interleukin-8 (IL-8), Tumor

15   Necrosis Factor-α (TNF-α.) and Tumor Necrosis Factor β (TNF-β.).

The term "antibody" as used herein refers to polyclonal and monoclonal antibodies and fragments thereof, and immunologic binding equivalents thereof. Antibodies may include, but are not limited to polyclonal antibodies, monoclonal antibodies (mAbs), humanized or chimeric antibodies, single chain antibodies, Fab

20   fragments, F(ab')$_2$ fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies, and epitope-binding fragments of any of the above.

The term "immunoglobulin polypeptide" as used herein refers to a constituent polypeptide of an antibody or a polypeptide derived therefrom. An "immunological polypeptide" may be, but is not limited to, an immunological heavy or light chain and

25   may include a variable region, a diversity region, joining region and a constant region or any combination, variant or truncated form thereof. The term "immunological polypeptides" further includes single-chain antibodies comprised of, but not limited to, an immunoglobulin heavy chain variable region, an immunoglobulin light chain variable region and optionally a peptide linker.

30   The terms "integrase" and "integrase activity" as used herein refer to a nucleic acid recombinase of the serine recombinase family of proteins.

The term "source of integrase activity" as used herein refers to a polypeptide or multimeric protein having serine recombinase (integrase) activity in an avian cell. The term may further refer to a polynucleotide encoding the serine recombinase, such as an mRNA, an expression vector, a gene or isolated gene that may be expressed as 5   the recombinase-specific polypeptide or protein.

The term "recombination site" as used herein refers to a polynucleotide stretch comprising a recombination site normally recognized and used by an integrase. For example, λ phage is a temperate bacteriophage that infects *E. coli*. The phage has one attachment site for recombination (attP) and the *E. coli* bacterial genome has an 10 attachment site for recombination (attB). Both of these sites are recombination sites for λ integrase. Recombination sites recognized by a particular integrase can be derived from a homologous system and associated with heterologous sequences, for example, the attP site can be placed in other systems to act as a substrate for the integrase.

15   The term "pseudo-recombination site" as used herein refers to a site at which an integrase can facilitate recombination even though the site may not have a sequence identical to the sequence of its wild-type recombination site. For example, a phiC31 integrase and vector carrying a phiC31 wild-type recombination site can be placed into an avian cell. The wild-type recombination sequence aligns itself with a sequence 20 in the avian cell genome and the integrase facilitates a recombination event. When the sequence from the genomic site in the avian cell, where the integration of the vector took place, is examined, the sequence at the genomic site typically has some identity to, but may not be identical with, the wild-type bacterial genome recombination site. The recombination site in the avian cell genome is considered to be a pseudo-25 recombination site (e.g., a pseudo-attP site) at least because the avian cell is heterologous to the normal phiC31 phage/bacterial cell system. The size of the pseudo-recombination site can be determined through the use of a variety of methods including, but not limited to, (i) sequence alignment comparisons, (ii) secondary structural comparisons, (iii) deletion or point mutation analysis to find the functional 30 limits of the pseudo-recombination site, and (iv) combinations of the foregoing.

A nucleic acid fragment of interest may be a trait-producing sequence, by which it is meant a sequence conferring a non-native trait upon the cell in which the protein encoded by the trait-producing sequence is expressed. The term "non-native" when used in the context of a trait-producing sequence means that the trait produced is

5   different than one would find in an unmodified organism which can mean that the organism produces high amounts of a natural substance in comparison to an unmodified organism, or produces a non-natural substance. For example, the genome of a bird could be modified to produce proteins not normally produced in birds such as, for instance, human or mouse antibodies, human cytokines, etc. Other useful traits

10   include disease resistance, meat flavor, animal size, and the like.

A nucleic acid fragment of interest may additionally be a "marker nucleic acid" or expressed as a "marker polypeptide". Marker genes encode proteins that can be easily detected in transformed cells and are, therefore, useful in the study of those cells. Examples of suitable marker genes include β-galactosidase, green or yellow

15   fluorescent proteins, enhanced green fluorescent protein, chloramphenicol acetyl transferase, luciferase, and the like. Such regions may also include those 5' noncoding sequences involved with initiation of transcription and translation, such as the enhancer, TATA box, capping sequence, CAAT sequence, and the like

The term "transformed" as used herein refers to a heritable alteration in a cell

20   resulting from the uptake of a heterologous DNA.

The term "trisomic" as used herein refers to a cell or animal, such as an avian cell or bird that has a 2n+1 chromosomal complement, where n is the haploid number of chromosomes, for the animal species concerned.

Techniques useful for isolating and characterizing the nucleic acids and

25   proteins of the present invention are well known to those of skill in the art and standard molecular biology and biochemical manuals may be consulted to select suitable protocols without undue experimentation. See, for example, Sambrook et al, 1989, "Molecular Cloning: A Laboratory Manual", 2nd ed., Cold Spring Harbor, the content of which is herein incorporated by reference in its entirety.

30

Abbreviations

Abbreviations used in the present specification include the following: aa, amino acid(s); bp, base pair(s); kb, kilobase; att, bacterial recombination attachment site; IU, infectious units.

5   In the standard method of integrase mediated-transgenesis, a serine recombinase integrase mediates recombination between an attB site on a transgene vector and a pseudo attP site on a chromosome. In the method of the invention for integrase-mediated transgenesis, a heterologous wild-type attP site can be integrated into an avian nuclear genome to create a transgenic cell line or bird. A serine

10   recombinase (integrase) and an attB-bearing transgene vector are then introduced into cells harboring the heterologous attP site, or into embryos derived from birds which bear the attP recombination site. The locations of attP and attB may be reversed such that the attB site is inserted into an avian chromosome and the attP sequence resides in an incoming transgene vector. In either case, the att site of the introduced vector

15   would then preferentially recombine with the integrated heterologous att site in the genome of the recipient cell.

The methods of the invention are based, in part, on the discovery that there exist in avian genomes a number of specific nucleic acid sequences, termed pseudo-recombination sites, the sequences of which may be distinct from wild-type

20   recombination sites but which can be recognized by a site-specific integrase and used to promote the efficient insertion of heterologous genes or polynucleotides into the targeted avian nuclear genome. The inventors have identified pseudo-recombination sites in avian cells capable of recombining with a recombination site, such as an attB site within a recombinant nucleic acid molecule introduced into the target avian cell.

25   The invention is also based on the prior integration of a heterologous att recombination site, typically isolated from a bacteriophage or a modification thereof, into the genome of the target avian cell.

Integration into a predicted chromosomal site is useful to improve the predictability of expression, which is particularly advantageous when creating

30   transgenic avians. Transgenesis by methods that result in insertion of the transgene

into random positions of the avian genome is unpredictable since the transgene may not express at the expected levels or in the predicted tissues.

The invention as disclosed herein, therefore, provides methods for site-specifically genetically transforming an avian nuclear genome. In general, an avian
5   cell having a first recombination site in the nuclear genome is transformed with a site-specific polynucleotide construct comprising a second recombination sequence and one or more polynucleotides of interest. Into the same cell, integrase activity is introduced that specifically recognizes the first and second recombination sites under conditions such that the polynucleotide sequence of interest is inserted into the nuclear
10   genome via an integrase-mediated recombination event between the first and second recombination sites.

The integrase activity, or a source thereof, can be introduced into the avian cell prior to, or concurrent with, the introduction of the site-specific construct. The integrase can be delivered to a cell as a polypeptide, or by expressing the integrase
15   from a source polynucleotide such as an mRNA or from an expression vector that encodes the integrase, either of which can be delivered to the target avian cell before, during or after delivery of the polynucleotide of interest. Any integrase that has activity in an avian cell may be useful in the present invention, including HK022 (Kolot *et al.*, *Biotechnol. Bioeng.*, 84: 56-60 (2003)). Preferably, the integrase is a
20   serine recombinase as described, for example, by Smith & Thorpe, in *Mol. Microbiol.*, 44: 299-307 (2002). More preferably, the integrase is a bacteriophage integrase such as, but not limited to, TP901-1 (Stoll *et al.*, *J. Bact.*, 184: 3657-3663 (2002); Olivares *et al.*, *Gene*, 278:167-176 (2001). Most preferably, the integrase is from the phage phiC31.

25   The nucleotide sequence of the junctions between an integrated transgene into the attP (or attB site) would be known. Thus, a PCR assay can be designed by one of skill in the art to detect when the integration event has occurred. The PCR assay for integration into a heterologous wild-type attB or attP site can also be readily incorporated into a quantitative PCR assay using TAQMAN™ or related technology so
30   that the efficiency of integration can be measured.

The minimal attB and attP sites able to catalyze recombination mediated by

the phiC31 integrase are 34 and 39 bp, respectively. In cell lines that harbor a heterologous integrated attP site, however, integrase has a preference for the inserted attP over any pseudo-attP sites of similar length, because pseudo-attP sites have very low sequence identity (between 10 to 50% identity) compared to the more efficient

5   wild-type attP sequence. It is within the scope of the methods of the invention, however, for the recombination site within the target avian genome to be a pseudo-att site such as a pseudo-attP site or an attP introduced into an avian genome.

The sites used for recognition and recombination of phage and bacterial DNAs (the native host system) are generally non-identical, although they typically have a

10   common core region of nucleic acids. The bacterial sequence is generally called the attB sequence (bacterial attachment) and the phage sequence is called the attP sequence (phage attachment). Because they are different sequences, recombination will result in a stretch of nucleic acids (called attL or attR for left and right) that is neither an attB sequence or an attP sequence, and likely is functionally unrecognizable

15   as a recombination site to the relevant enzyme, thus removing the possibility that the enzyme will catalyze a second recombination reaction that would reverse the first.

The integrase may recognize a recombination site where sequence of the 5' region of the recombination site can differ from the sequence of the 3' region of the recombination sequence. For example, for the phage phiC31 attP (the phage

20   attachment site), the core region is 5'-TTG-3' the flanking sequences on either side are represented here as attP5' and attP3', the structure of the attP recombination site is, accordingly, attP5'-TTG-attP3'. Correspondingly, for the native bacterial genomic target site (attB) the core region is 5'-TTG-3', and the flanking sequences on either side are represented here as attB5' and attB3', the structure of the attB recombination

25   site is, accordingly, attB5'-TTG-attB3'. After a single-site, phiC31 integrase-mediated recombination event takes place between the phiC31 phage and the bacterial genome, the result is the following recombination product: attB5'-TTG-attP3'{phiC31 vector sequences}attP5'-TTG-attB3'. In the method of invention, the attB site will be within a recombinant nucleic acid molecule that may be delivered to a target avian cell. The

30   corresponding attP (or pseudo-attP) site will be within the avian cell nuclear genome. Consequently, after phiC31 integrase mediated recombination, the recombination

product, the nuclear genome with the integrated heterologous polynucleotide will have the sequence attP5'-TTG-attB3'{heterologous polynucleotide}-attB5'-TTG-attP3'. Typically, after recombination the post-recombination recombination sites are no longer able to act as substrate for the phiC31 integrase. This results in stable

5   integration with little or no integrase mediated excision.

While the preferred recombination site to be included in the recombinant nucleic acid molecules and modified chromosomes of the present invention is the attP site, it is contemplated that any attP-like site may be used if compatible with the attB site. For instance, any pseudo-attP site of the chicken genome may be identified

10   according to the methods of Example 7 below and used as a heterologous att recombination site. Such attP-like sites may have a sequence that is at least 25% identical to SEQ ID NO: 11 as shown in Fig. 19, such as described in Groth *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 97: 5995-6000 (2000) incorporated herein by reference in its entirety. Preferably the selected site will have at least the same degree of

15   efficiency of recombination as the attP site (SEQ ID NO: 11) itself.

In the methods of the present invention, the recipient avian cell population may be an isolated avian cell line such as, for example, DF-1 chicken fibroblasts, chicken DT40 cells or a cell population derived from an early stage embryo such as a chicken stage I or stage X embryo. A particularly useful avian cell population is

20   blastodermal cells isolated from an early stage I embryo or a stage X avian embryo. The methods of the present invention, therefore, include steps for the isolation of blastodermal cells that are then suspended in a cell culture medium or buffer for maintaining the cells in a viable state, and which allows the cell suspension to contact the nucleic acids of the present invention. It is also within the scope of the invention

25   for the nucleic acid construct and the source of integrase activity to be delivered directly to an avian embryo such as a blastodermal layer, or to a tissue layer of an adult bird such as the lining of an oviduct.

When the recipient avian cell population is isolated from an early stage avian embryo, the embryos must first be isolated. For stage I avian embryos from, for

30   example, a chicken, a fertilized ovum is surgically removed from a bird before the deposition of the outer hard shell has occurred. The nucleic acids for integrating a

heterologous nucleic acid into a recipient avian cell genome may then be delivered to isolated embryos by lipofection, microinjection (as described in Example 6 below) or electroporation and the like. After delivery of the nucleic acid, the transfected embryo and its yolk may be deposited into the infundibulum of a recipient hen for the

5   deposition of egg white proteins and a hard shell, and laying of the egg. Stage X avian embryos are obtained from freshly laid fertilized eggs and the blastodermal cells isolated as a suspension of cells in a medium, as described in Example 4 below. Isolated stage X blastodermal cell populations, once transfected, may be injected into recipient stage X embryos and the hard shell eggs resealed according to the methods

10   described in U.S. Patent No. 6,397,777.

In the methods of the invention, once a heterologous nucleic acid is delivered to the recipient avian cell, the integrase activity is expressed. The expressed integrase (or injected integrase polypeptide) then mediates recombination between the att site of the heterologous nucleic acid molecule, and the att (or pseudo att) site within the

15   genomic DNA of the recipient avian cell.

It is within the scope of the present invention for the integrase-encoding sequence and a promoter operably linked thereto to be included in the delivered nucleic acid molecule and that expression of the integrase activity occurs before integration of the heterologous nucleic acid into the avian cell genome. Preferably,

20   the integrase-encoding nucleic acid sequence and associated promoter are in an expression vector that may be co-delivered to the recipient avian cell with the heterologous nucleic acid molecule to be integrated into the recipient genome.

One suitable integrase expressing expression vector for use in the present invention is pCMV-C31int (SEQ ID NO: 1) as shown in Fig. 9, and described in

25   Groth et al., Proc. Natl. Acad. Sci. U.S.A. 97: 5995-6000 (2000), incorporated herein by reference in its entirety. In pCMV-C31int, expression of the integrase-encoding sequence is driven by the CMV promoter. However, any promoter may be used that will give expression of the integrase in a recipient avian cell, including operably linked avian-specific gene expression control regions of the avian ovalbumin,

30   lysozyme, ovomucin, ovomucoid gene loci, viral gene promoters, inducible promoters, the RSV promoter and the like.

The recombinant nucleic acid molecules of the present invention for delivery of a heterologous polynucleotide to the genome of a recipient avian cell may comprise a nucleotide sequence encoding the attB attachment site of *Streptomyces ambofaciens* as described in Thorpe & Smith, *Proc. Natl. Acad. Sci. U.S.A.* 95: 5505-5510 (1998).

5   The nucleic acid molecule of the present invention further comprises an expression cassette for the expression in a recipient avian cell of a heterologous nucleic acid encoding a desired heterologous polypeptide. Optionally, the nucleic acid molecules may further comprise a marker such as, but not limited to, a puromycin resistance gene, a luciferase gene, EGFP, and the like.

10 It is contemplated that the expression cassette for introducing a desired heterologous polypeptide comprises a promoter operably linked to a nucleic acid encoding the desired polypeptide and, optionally, a polyadenylation signal sequence. Exemplary nucleic acids suitable for use in the present invention are more fully described in the examples below.

15 In the methods of the present invention, following delivery of the nucleic acid molecule and a source of integrase activity into an avian cell population, the cells are maintained under culture conditions suitable for the expression of the integrase and/or for the integrase to mediate recombination between the recombination site of the nucleic acid and recombination site in the genome of the recipient avian cell. When

20 the recipient avian cell is cultured *in vitro,* such cells may be incubated at 37° Celsius if the cells are chicken early stage blastodermal cells. They may then be injected into an embryo within a hard shell, which is resealed for incubation until hatching. Alternatively, the transfected cells may be maintained in *in vitro* culture.

25 *Site-Specific Nucleic Acid Constructs and Methods of Delivery to an Avian Cell*

The present invention provides methods for the site-specific insertion of a heterologous nucleic acid molecule into the nuclear genome of an avian cell by delivering to a target avian cell that has a recombination site in its nuclear genome, a source of integrase activity, a site-specific construct that has another recombination

30 site and a polynucleotide of interest, and allowing the integrase activity to facilitate a recombination event between the two recombination sites, thereby integrating the

polynucleotide of interest into the avian nuclear genome.

(a) Expression vector nucleic acid molecules: A variety of recombinant nucleic acid expression vectors are suitable for use in the practice of the present invention. The site-specific constructs described herein can be constructed utilizing methodologies well known in the art of molecular biology (see, for example, *Ausubel* or *Maniatis*) in view of the teachings of the specification. As described above, the constructs are assembled by inserting into a suitable vector backbone a recombination site such as an attP or an attB site, a polynucleotide of interest operably linked to a gene expression control region of interest and, optionally a sequence encoding a positive selection marker. Polynucleotides of interest can include, but are not limited to, expression cassettes encoding a polypeptide to be expressed in the transformed avian cell or in a transgenic bird derived therefrom. The site-specific constructs are typically circular and may also contain selectable markers, an origin of replication, and other elements.

Any of the vectors of the present invention may also optionally include a sequence encoding a signal peptide that directs secretion of the polypeptide expressed by the vector from the transgenic cells, for instance, from tubular gland cells of the oviduct. This aspect of the invention effectively broadens the spectrum of exogenous proteins that may be deposited in the whites of avian eggs using the methods of the invention. Where an exogenous polypeptide would not otherwise be secreted, the vector bearing the coding sequence can be modified to comprise, for instance, about 60 bp encoding a signal peptide. The DNA sequence encoding the signal peptide is inserted in the vector such that the signal peptide is located at the N-terminus of the polypeptide encoded by the vector.

The expression vectors of the present invention can comprise an avian transcriptional regulatory region for directing expression of either fusion or non-fusion proteins. With fusion vectors, a number of amino acids are usually added to the desired expressed target gene sequence such as, but not limited to, a polypeptide sequence for thioredoxin. A proteolytic cleavage site may further be introduced at a site between the target recombinant protein and the fusion sequence. Additionally, a region of amino acids such as a polymeric histidine region may be introduced to allow binding of the fusion protein to metallic ions such as nickel bonded to a solid support,

for purification of the fusion protein. Once the fusion protein has been purified, the cleavage site allows the target recombinant protein to be separated from the fusion sequence. Enzymes suitable for use in cleaving the proteolytic cleavage site include, but are not limited to, Factor Xa and thrombin. Fusion expression vectors that may be

5   useful in the present invention include pGex (Amrad Corp., Melbourne, Australia), pRIT5 (Pharmacia, Piscataway, NJ) and pMAL (New England Biolabs, Beverly, MA), that fuse glutathione S-transferase, protein A, or maltose E binding protein, respectively, to a desired target recombinant protein.

Epitope tags are short peptide sequences that are recognized by epitope

10   specific antibodies. A fusion protein comprising a recombinant protein and an epitope tag can be simply and easily purified using an antibody bound to a chromatography resin, for example.   The presence of the epitope tag furthermore allows the recombinant protein to be detected in subsequent assays, such as Western blots, without having to produce an antibody specific for the recombinant protein itself.

15   Examples of commonly used epitope tags include V5, glutathione-S-transferase (GST), hemaglutinin (HA), the peptide Phe-His-His-Thr-Thr, chitin binding domain, and the like.

Preferred gene expression control regions for use in avian cells include, but are not limited to, avian specific promoters such as the chicken lysozyme, ovalbumin, or

20   ovomucoid promoters, and the like. Particularly useful are tissue-specific promoters such as avian oviduct promoters that allow for expression and delivery of a heterologous polypeptide to an egg white.

Viral promoters serve the same function as bacterial or eukaryotic promoters and either provide a specific RNA polymerase in trans (bacteriophage T7) or recruit

25   cellular factors and RNA polymerase (SV40, RSV, CMV). Viral promoters may be preferred as they are generally particularly strong promoters. A preferred promoter for use in avian cells is the RSV promoter.

Selection markers are valuable elements in expression vectors as they provide a means to select for growth of only those cells that contain a vector. Common

30   selectable marker genes include those for resistance to antibiotics such as ampicillin, puromycin, tetracycline, kanamycin, bleomycin, streptomycin, hygromycin, neomycin,

ZEOCIN™, and the like.

Another element useful in an expression vector is an origin of replication. Replication origins are unique DNA segments that contain multiple short repeated sequences that are recognized by multimeric origin-binding proteins and that play a

5   key role in assembling DNA replication enzymes at the origin site. Suitable origins of replication for use in expression vectors employed herein include *E. coli* oriC, colE1 plasmid origin, and the like.

A further useful element in an expression vector is a multiple cloning site or polylinker. Synthetic DNA encoding a series of restriction endonuclease recognition

10   sites is inserted into a vector, for example, downstream of the promoter element. These sites are engineered for convenient cloning of DNA into the vector at a specific position.

Elements such as the foregoing can be combined to produce expression vectors suitable for use in the methods of the invention. Those of skill in the art will be able

15   to select and combine the elements suitable for use in their particular system in view of the teachings of the present specification.

(b) Genetically modified avian and artificial chromosomes: The present invention further provides modified chromosomes, either isolated avian or artificial chromosomes, are useful vectors to shuttle transgenes or gene clusters into the avian

20   genome. By delivering the modified or artificial chromosome to an isolated recipient cell, the target cell, and progeny thereof, become trisomic. Preferably, an additional or triosomic chromosome will not affect the subsequent development of the recipient cell and/or an embryo, nor interfere with the reproductive capacity of an adult bird developed from such cells or embryos. The chromosome also should be stable within

25   chicken cells. An effective method is also required to isolate a population of chromosomes for delivery into chicken embryos or early cells.

A number of artificial chromosomes are useful in the methods of the invention, including, for instance, a human chromosome modified to work as an artificial chromosome in a heterologous species as described, for example, for mice

30   (Tomizuka *et al., Proc. Natl. Acad. Sci. U.S.A.* 97: 722-727 (2000); for cattle (Kuroiwa *et al.,* Nat. *Biotechnol.* 20: 889-894 (2002); a mammalian artificial

chromosome used in mice (Co *et al.*, *Chromosome Res.* 8: 183-191 (2000), or in viable triploid chickens (Thorne *et al.*, *Cytogenet. Cell Genet.* 57: 206-210 (1991) and Thorne, *et al.*, *J. Hered.* 88: 495-498 (1997). Chickens that are trisomic for microchromosome 16 have been described (Miller *et al.*, *Proc. Natl. Acad. Sci. U.S.A.*

5   93: 3958-3962 (1996); Muscarella *et al.*, *J. Cell Biol.* 101: 1749-1756 (1985). In these cases, triploidy and trisomy occurred naturally, and illustrate that an extra copy of one or more of the chicken chromosomes is compatible with normal development and reproductive capacity.

A useful chromosome isolation protocol can comprise the steps of inserting a

10   lac-operator sequence (Robinett *et al. J. Cell Biol.* 135: 1685-1700 (1996) into an isolated chromosome and, optionally, inserting a desired transgene sequence within the same chromosome. Preferably, the lac operator region is a concatamer of a plurality of lac operators for the binding of multiple lac repressor molecules. Insertion can be accomplished, for instance, by identifying a region of known nucleotide

15   sequence associated with a particular avian chromosome. A recombinant DNA molecule may be constructed that comprises the identified region, a recombination site such as attB or attP and a lac-operator concatamer. The recombinant molecule is delivered to an isolated avian cell, preferably, but not limited to, chicken DT40 cells that have elevated homologous recombination activity compared to other avian cell

20   lines, whereupon homologous recombination will integrate the heterologous recombination site and the lac-operator concatamer into the targeted chromosome as shown in the schema illustrated in Fig. 20. A tag-polypeptide comprising a label domain and a lac repressor domain is also delivered to the cell, preferably by expression from a suitable expression vector. The nucleotide sequence coding for a

25   GFP-lac-repressor fusion protein (Robinett *et al.*, *J. Cell Biol.* 135: 1685-1700 (1996)) may be inserted into the same chromosome as the lac-operator insert. The lac repressor sequence, however, can also be within a different chromosome. An inducible promoter may also be used to allow the expression of the GFP-lac-repressor only after chromosome is to be isolated.

30   Induced expression of the GPF-lac-repressor fusion protein will result in specific binding of the tag fusion polypeptide to the lac-operator sequence for

identification and isolation of the genetically modified chromosome. The tagged mitotic chromosome can be isolated using, for instance, flow cytometry as described in de Jong *et al. Cytometry* 35: 129-133 (1999) and Griffin *et al.* Cytogenet. *Cell Genet.* 87: 278-281 (1999).

5   A tagged chromosome can also be isolated using microcell technology requiring treatment of cells with the mitotic inhibitor colcemid to induce the formation of micronuclei containing intact isolated chromosomes within the cell. Final separation of the micronuclei is then accomplished by centrifugation in cytochalasin as described by Killary & Fournier in Methods Enzymol. 254: 133-152

10   (1995). Further purification of microcells containing only the desired tagged chromosome could be done by flow cytometry. It is contemplated, however, that alternative methods to isolate the mitotic chromosomes or microcells, including mechanical isolation or the use of laser scissors and tweezers, and the like.

15   *Delivery of a Site-Specific Nucleic Acid to a Recipient Avian Cell or Embryo.*
       (a) Delivery of polynucleotide constructs.

       Most non-viral methods of gene transfer rely on normal mechanisms used by eukaryotic cells for the uptake and intracellular transport of macromolecules. In preferred embodiments, non-viral gene delivery systems of the present invention rely

20   on endocytic pathways for the uptake of the subject transcriptional regulatory region and operably linked polypeptide-encoding nucleic acid by the targeted cell. Exemplary gene delivery systems of this type include liposomal derived systems, poly-lysine conjugates, and artificial viral envelopes. Modified chromosomes as described above may be delivered to isolated avian embryonic ells for subsequent

25   introduction to an embryo.

       In a representative embodiment, a nucleic acid molecule can be entrapped in liposomes bearing positive charges on their surface (*e.g.*, lipofectins) and (optionally) which are tagged with antibodies against cell surface antigens of the target tissue (Mizuno *et al.*, 1992, *NO Shinkei Geka* 20: 547-551; PCT publication WO91/06309;

30   Japanese patent application 1047381; and European patent publication EP-A-43075, all of which are incorporated herein by reference in their entireties).

In similar fashion, the gene delivery system can comprise an antibody or cell surface ligand that is cross-linked with a gene binding agent such as polylysine (see, for example, PCT publications WO93/04701, WO92/22635, WO92/20316, WO92/19749, and WO92/06180, all of which are incorporated herein by reference in

5   their entireties). It will also be appreciated that effective delivery of the subject nucleic acid constructs via receptor-mediated endocytosis can be improved using agents which enhance escape of genes from the endosomal structures. For instance, whole adenovirus or fusogenic peptides of the influenza HA gene product can be used as part of the delivery system to induce efficient disruption of DNA-containing

10   endosomes (Mulligan *et al.*, 1993, *Science* 260-926; Wagner *et al.*, 1992, *Proc. Natl. Acad. Sci.* 89:7934-7938; and Christiano *et al.*, 1993, *Proc. Natl. Acad. Sci.* 90:2122-2126, all of which are incorporated herein by reference in their entireties). It is further contemplated that a recombinant nucleic acid molecule of the present invention may be delivered to a target host cell by other non-viral methods including by gene gun,

15   microinjection, sperm-mediated transfer, or the like.

In yet another embodiment of the invention, an expression vector that comprises a heterologous attB recombination site and a region encoding a polypeptide deposited into an egg white are delivered to oviduct cells by in vivo electroporation. In this method, the luminal surface of an avian oviduct is surgically exposed. A

20   buffered solution of the expression vector and a source of integrase activity such as a second expression vector expressing integrase (for example pCMV-int) is deposited on the luminal surface. Electroporation electrodes are then positioned on either side of the oviduct wall, the luminal electrode contacting the expression vector solution. After electroporation, the surgical incisions are closed. The electroporation will

25   deliver the expression vectors to some, if not all, treated recipient oviduct cells to create a tissue-specific chimeric animal. Expression of the integrase allows for the integration of the heterologous polynucleotide into the nuclear genomes of recipient oviduct cells. While this method may be used with any bird, a preferred recipient is a chicken due to the size of the oviduct. More preferred is a transgenic bird that has a

30   transgenic attP recombinant site in the nuclear genomes of recipient oviduct cells, thus increasing the efficiency of integration of the expression vector.

The attB/P integrase system is preferred in the in vivo electroporation method to allow the formation of stable genetically transformed oviduct cells that otherwise progressively lose the heterologous expression vector.

5   The stably modified oviduct cells will express the heterologous polynucleotide and deposit the resulting polypeptide into the egg white of a laid egg. For this purpose, the expression vector will further comprise an oviduct-specific promoter such as ovalbumin or ovomucoid operably linked to the desired heterologous polynucleotide.

(b) Delivery of chromosomes to avian cells.

10   Another aspect of the invention is the generation of a trisomic avian cell comprising a genetically modified extra chromosome. The extra chromosome may be an artificial chromosome or an isolated avian chromosome that has been genetically modified. Introduction of the extra chromosome to an avian cell will generate a trisomic cell with 2n+1 chromosomes, where n is the haploid number of chromosomes

15   of a normal avian cell.

Delivery of an isolated chromosome into an isolated avian cell or embryo can be accomplished in several ways. Isolated mitotic chromosomes or a micronucleus containing an interphase chromosome can be injected into early stage I embryos by cytoplasmic injection. The injected zygote would then be surgically transferred to a

20   recipient hen for the production and laying of a hard shell egg. This hard shell egg would then be incubated until hatching of a chick.

Isolated microcells can be fused to primordial germ cells (PGCs) isolated from the blood stream of late stage 15 embryos as described by Killary & Fournier in *Methods Enzymol.* 254: 133-152 (1995). The PGC/microcell hybrids can then be

25   transplanted into the blood stream of a recipient embryo to produce germline chimeric chickens. (See Naito *et al., Mol. Reprod. Dev.* 39: 153-161 (1994)). The manipulated eggs would then incubated until hatching of the bird.

Blastodermal cells isolated from stage X embryos can be transfected with isolated mitotic chromosomes. Following *in vitro* transfection, the cells are

30   transplanted back into stage X embryos as described, for example, in Etches *et al., Poult. Sci.,* 72: 882-829 (1993), and the manipulated eggs are incubated to hatching.

Stage X blastodermal cells can also be fused with isolated microcells and then transplanted back into to stage X embryos or fused to somatic cells to be used as nuclear donors for nuclear transfer as described by Kuroiwa *et al.*, *Nat. Biotechnol.* 20: 889-894 (2002).

5   Chromosomal vectors, as described above, may be delivered to a recipient avian cell by, for example, microinjection, liposomal delivery or microcell fusion.

Delivering a Source of Integrase Activity to an Avian Cell

In the methods of the invention, a site-specific integrase is introduced into an
10   avian cell whose genome is to be modified. Methods of introducing functional proteins into cells are well known in the art. Introduction of purified integrase protein can ensure a transient presence of the protein and its activity. Thus, the lack of permanence associated with most expression vectors is not expected to be detrimental.

The integrase used in the practice of the present invention can be introduced
15   into a target cell before, concurrently with, or after the introduction of a site-specific vector. The integrase can be directly introduced into a cell as a protein, for example, by using liposomes, coated particles, or microinjection, or into the blastodermal layer of an early stage avian embryo by microinjection. A source of the integrase can also be delivered to an avian cell by introducing to the cell an mRNA encoding the
20   integrase and which can be expressed in the recipient cell as an integrase polypeptide. Alternately, a DNA molecule encoding the integrase can be introduced into the cell using a suitable expression vector.

The present invention provides novel nucleic acid vectors and methods of use that allow the phiC31 integrase to efficiently integrate a heterologous nucleic acid into
25   an avian genome. A novel finding is that the phiC31 integrase is remarkably efficient in avian cells and increases the rate of integration of heterologous nucleic acid at least 30-fold over that of random integration. Furthermore, the phiC31 integrase works equally well at 37°C and 41°C, indicating that it will function in the environment of the developing avian embryo, as shown in Example 1.

30   The site-specific vector components described above are useful in the construction of expression cassettes containing sequences encoding an integrase. One

integrase-expressing vector useful in the methods of the invention is pCMV-C31int (SEQ ID NO: 1 as shown in Fig. 9) where the phiC31 integrase is encoded by a region under the expression control of the strong CMV promoter. Another preferred promoter generally useful in avian cells is the RSV promoter as used in SEQ ID NO: 9

5   shown in Fig. 17. Expression of the integrase is typically desired to be transient. Accordingly, vectors providing transient expression of the integrase are preferred. However, expression of the integrase can be regulated in other ways, for example, by placing the expression of the integrase under the control of a regulatable promoter (i.e., a promoter whose expression can be selectively induced or repressed).

10

Transgenic Avian Cells.

Another aspect of the present invention is an avian cell genetically modified with a transgene vector according to the present invention and described above. For example, in one embodiment, the transformed cell can be a chicken early stage

15   blastodermal cell or a genetically transformed cell line, including a sustainable cell line. The transfected cell according to the present invention may comprise a transgene stably integrated into the nuclear genome of the recipient cell, thereby replicating with the cell so that each progeny cell receives a copy of the transfected nucleic acid. A particularly useful cell line for the delivery and integration of a transgene comprises a

20   heterologous attP site that can increase the efficiency of integration of a polynucleotide by phiC31 integrase and, optionally, a region for expressing the integrase.

A retroviral vector can be used to deliver the att site into the avian genome since an attP or attB site is less than 300 bp. For example, the attP site can be inserted

25   into the NLB retroviral vector, which is based on the avian leukosis virus genome. A lentiviral vector is a particularly suitable vector because lentiviral vectors can transduce non-dividing cells, so that a higher percentage of cells will have an integrated attP site.

The lacZ region of NLB is replaced by the attP sequence. A producer cell line

30   would be created by transformation of, for example, the Isolde cell line capable of producing a packaged recombinant NLB-attP virus pseudo-typed with the envA

envelope protein. Supernatant from the Isolde NLB-attP line is concentrated by centrifugation to produce high titer preparations of the retroviral vector that can then be used to deliver the attP site to the genome of an avian cell, as described in Example 9 below.

5   An attP-containing line of transgenic birds are a source of attP transgenic embryos and embryonic cells. Fertile zygotes and oocytes bearing a heterologous attP site in either the maternal, paternal, or both, genomes can be used for transgenic insertion of a desired heterologous polynucleotide. A transgene vector bearing an attB site, for example, would be injected into the cytoplasm along with either an integrase

10   expression plasmid, mRNA encoding the integrase or the purified integrase protein. The oocyte or zygote is then cultured to hatch by *ex ovo* methods or reintroduced into a recipient hen such that the hen lays a hard shell egg the next day containing the injected egg.

          In another example, fertile stage VII-XII embryos hemizygous or homozygous

15   for the heterologous attP sequence, are used as a source of blastodermal cells. The cells are harvested and then transfected with a transgene vector bearing an attB site along with a source of integrase. The transfected cells are then injected into the subgerminal cavity of windowed fertile eggs. The chicks that hatch will bear the transgene integrated into the attP site in a percentage of their somatic and germ cells.

20   To obtain fully transgenic birds, chicks are raised to sexual maturity and those that are positive for the transgene in their semen are bred to non-transgenic mates.

          In various embodiments, the genetically engineered cells of the invention may contain an integrase specifically recognizing recombination sites and which is introduced into genetically engineered cells containing a nucleic acid construct of the

25   invention under conditions such that the nucleic acid sequence(s) of interest will be inserted into the nuclear genome. Methods for introducing such an integrase into a cell are described above.

          In some embodiments, the site-specific integrase is introduced into the cell as a polypeptide. In alternative embodiments, the site-specific integrase is introduced into

30   the transgenic cell as a polynucleotide encoding the integrase, such as an expression

cassette optionally carried on a transient expression vector, and comprising a polynucleotide encoding the recombinase.

In one embodiment, the invention is directed to methods of using a vector for site-specific integration of a heterologous nucleotide sequence into the genome of an avian cell, the vector comprising a circular backbone vector, a polynucleotide of interest operably linked to a promoter, and a first recombination site, wherein the genome of the cell comprises a second recombination site and recombination between the first and second recombination sites is facilitated by phiC31 integrase. In certain embodiments, the integrase facilitates recombination between a bacterial genomic recombination site (attB) and a phage genomic recombination site (attP).

In another embodiment, the invention is directed to an avian cell having a transformed genome comprising an integrated heterologous polynucleotide of interest whose integration, mediated by phiC31 integrase, was into a recombination site native to the avian cell genome and the integration created a recombination-product site comprising the polynucleotide sequence. In yet another embodiment, integration of the polynucleotide was into a recombination site not native to the avian cell genome, but instead into a heterologous recombination site engineered into the avian cell genome.

In further embodiments, the invention is directed to transgenic birds comprising a modified cell and progeny thereof as described above, as well as methods of producing the same.

Cells genetically modified to carry a heterologous attB or attP site by the methods of the present invention can be maintained under conditions that, for example, keep them alive but do not promote growth, promote growth of the cells, and/or cause the cells to differentiate or dedifferentiate. Cell culture conditions may be permissive for the action of the integrase in the cells, although regulation of the activity of the integrase may also be modulated by culture conditions (e.g., raising or lowering the temperature at which the cells are cultured).

One aspect of the invention is a method for generating a genetically modified avian cell, and progeny thereof, using a tagged chromosome, the method comprising the steps of providing an isolated modified chromosome comprising a lac operator

region and a first recombination site, delivering the modified chromosome to a avian cell, thereby generating a trisomic avian cell, delivering to the avian cell a source of a tagged polypeptide comprising a fluorescent domain and a lac repressor domain, delivering a source of integrase activity to the avian cell, delivering a polynucleotide comprising a second recombination site and a region encoding a polypeptide to the avian cell, maintaining the avian cell under conditions suitable for the integrase to mediate recombination between the first and second recombination sites, thereby integrating the polynucleotide into the modified chromosome and generating a genetically modified avian cell, expressing the tag polypeptide by the avian cell, allowing the tag polypeptide to bind to the modified chromosome so as to label the modified chromosome, and isolating the modified chromosome by selecting modified chromosomes having a tag polypeptide bound thereto.

In one embodiment of the invention, the second avian cell is selected from the group consisting of a stage VII-XII blastodermal cell, a stage I embryo, a stage X embryo; an isolated primordial germ cell, an isolated non-embryonic cell, and an oviduct cell.

In various embodiments, the isolated modified chromosome is an avian chromosome or an artificial chromosome.

In other embodiments of the invention, the step of providing an isolated modified chromosome comprising a lac operator region and a first recombination site comprises the steps of generating a trisomic avian cell by delivering to an isolated avian cell an isolated chromosome and a polynucleotide comprising a lac operator and a second recombination site, maintaining the trisomic cell under conditions whereby the heterologous polynucleotide is integrated into the chromosome by homologous recombination, delivering to the avian cell a source of a tag polypeptide to label the chromosome, and isolating the labeled chromosome.

In one embodiment of the invention, the lac operator region is a concatamer of lac operators. In other embodiments of the invention, the tag polypeptide is expressed from an expression vector.

In one embodiment of the invention, the tag polypeptide is microinjected into the cell. In various embodiments of the invention, the method of delivery of a

chromosome to an avian cell is selected from the group consisting of liposome delivery, microinjection, microcell, electroporation and gene gun delivery, or a combination thereof.

In embodiments of the invention, the fluorescent domain of the tag polypeptide is GFP.

In another embodiment of the invention, the method further comprises the step of delivering the second avian cell to an avian embryo. The embryo may be maintained under conditions suitable for hatching as a chick.

In one embodiment of the invention, the second avian cell is maintained under conditions suitable for the proliferation of the cell, and progeny thereof.

In various embodiments of the invention, the source of integrase activity is delivered to a first avian cell as a polypeptide or expressed from a polynucleotide, said polynucleotide being selected from an mRNA and an expression vector.

In one embodiment of the invention, the tag polypeptide activity is delivered to the avian cell as a polypeptide or expressed from a polynucleotide operably linked to a promoter. In another embodiment of the invention, the promoter is an inducible promoter. In yet another embodiment of the invention, the integrase is phiC31 integrase and in various embodiments of the invention, the first and second recombination sites are selected from an attB and an attP site, but wherein the first and second sites are not identical.

Expression of Heterologous Proteins by Site-Specific Genetic Transformation of Avian Cells.

Another aspect of the present invention is a method of expressing a heterologous polypeptide in an avian cell by stably transfecting a cell by using site-specific integrase-mediation and a recombinant nucleic acid molecule, as described above, and culturing the transfected cell under conditions suitable for expression of the heterologous polypeptide under the control of the avian transcriptional regulatory region.

The protein of the present invention may be produced in purified form by any known conventional techniques. For example, chicken cells, an egg or an egg white

may be homogenized and centrifuged. The supernatant may then be subjected to sequential ammonium sulfate precipitation and heat treatment. The fraction containing the protein of the present invention is subjected to gel filtration in an appropriately sized dextran or polyacrylamide column to separate the proteins. If

5   necessary, the protein fraction may be further purified by HPLC or other methods well known in the art of protein purification.

The methods of the invention are useful for expressing nucleic acid sequences that are optimized for expression in avian cells and which encode desired polypeptides or derivatives and fragments thereof. Derivatives include, for instance, polypeptides

10   with conservative amino acid replacements, that is, those within a family of amino acids that are related in their side chains (commonly known as acidic, basic, nonpolar, and uncharged polar amino acids). Phenylalanine, tryptophan, and tyrosine are sometimes classified jointly as aromatic amino acids and other groupings are known in the art (see, for example, "Biochemistry", 2nd ed, L. Stryer, ed., W.H. Freeman &

15   Co.,1981). Peptides in which more than one replacement has taken place can readily be tested for activity in the same manner as derivatives with a single replacement, using conventional polypeptide activity assays (e.g. for enzymatic or ligand binding activities).

Regarding codon optimization, if the recombinant nucleic acid molecules are

20   transfected into a recipient chicken cell, the sequence of the nucleic acid insert to be expressed can be optimized for chicken codon usage. This may be determined from the codon usage of at least one, and preferably more than one, protein expressed in a chicken cell according to well known principles. For example, in the chicken the codon usage could be determined from the nucleic acid sequences encoding the

25   proteins such as lysozyme, ovalbumin, ovomucin and ovotransferrin of chicken. Optimization of the sequence for codon usage can elevate the level of translation in avian eggs.

The present invention provides methods for the production of a protein by an avian cell comprising the steps of maintaining an avian cell, transfecting with a first

30   expression vector and, optionally, a second expression vector, under conditions suitable for proliferation and/or gene expression and such that an integrase will

mediate site specific recombination at att sites. The expression vectors may each have a transcription unit comprising a nucleotide sequence encoding a heterologous polypeptide, wherein one polypeptide is an integrase, a transcription promoter, and a transcriptional terminator. The cells may then be maintained under conditions for the expression and production of the desired heterologous polypeptide(s).

The present invention further relates to methods for gene expression by avian cells from nucleic acid vectors, and transgenes derived therefrom, that include more than one polypeptide-encoding region wherein, for example, a first polypeptide-encoding region can be operatively linked to an avian promoter and a second polypeptide-encoding region is operatively linked to an Internal Ribosome Entry Sequence (IRES). It is contemplated that the first polypeptide-encoding region, the IRES and the second polypeptide-encoding region of a recombinant DNA of the present invention may be arranged linearly, with the IRES operably positioned immediately 5' of the second polypeptide-encoding region. This nucleic acid construct, when inserted into the genome of an avian cell or a bird and expressed therein, will generate individual polypeptides that may be post-translationally modified and combined in the white of a hard shell bird egg. Alternatively, the expressed polypeptides may be isolated from an avian egg and combined *in vitro*.

The invention, therefore, includes methods for producing multimeric proteins including immunoglobulins, such as antibodies, and antigen binding fragments thereof. Thus, in one embodiment of the present invention, the multimeric protein is an immunoglobulin, wherein the first and second heterologous polypeptides are immunoglobulin heavy and light chains respectively. Illustrative examples of this and other aspects of the present invention for the production of heterologous multimeric polypeptides in avian cells are fully disclosed in U.S. Patent Application No. 09/877,374, filed June 8, 2001, by *Rapp*, published as US-2002-0108132-Al on August 8, 2002, and U.S. Patent Application No. 10/251,364, filed September 18, 2002, by *Rapp*, both of which are incorporated herein by reference in their entirety.

Accordingly, the invention further provides immunoglobulin and other multimeric proteins that have been produced by transgenic avians of the invention.

In various embodiments, an immunoglobulin polypeptide encoded by the transcriptional unit of at least one expression vector may be an immunoglobulin heavy chain polypeptide comprising a variable region or a variant thereof, and may further comprise a D region, a J region, a C region, or a combination thereof. An immunoglobulin polypeptide encoded by an expression vector may also be an immunoglobulin light chain polypeptide comprising a variable region or a variant thereof, and may further comprise a J region and a C region. The present invention also contemplates multiple immunoglobulin regions that are derived from the same animal species, or a mixture of species including, but not only, human, mouse, rat, rabbit and chicken. In preferred embodiments, the antibodies are human or humanized.

In other embodiments, the immunoglobulin polypeptide encoded by at least one expression vector comprises an immunoglobulin heavy chain variable region, an immunoglobulin light chain variable region, and a linker peptide thereby forming a single-chain antibody capable of selectively binding an antigen.

Examples of therapeutic antibodies that can be used in methods of the invention include but are not limited to HERCEPTIN™ (Trastuzumab) (Genentech, CA) which is a humanized anti-HER2 monoclonal antibody for the treatment of patients with metastatic breast cancer; REOPRO™ (abciximab) (Centocor) which is an anti-glycoprotein IIb/IIIa receptor on the platelets for the prevention of clot formation; ZENAPAX™ (daclizumab) (Roche Pharmaceuticals, Switzerland) which is an immunosuppressive, humanized anti-CD25 monoclonal antibody for the prevention of acute renal allograft rejection; PANOREX™ which is a murine anti-17-IA cell surface antigen IgG2a antibody (Glaxo Wellcome/Centocor); BEC2 which is a murine anti-idiotype (GD3 epitope) IgG antibody (ImClone System); IMC-C225 which is a chimeric anti-EGFR IgG antibody (ImClone System); VITAXIN™ which is a humanized anti-αVβ3 integrin antibody (Applied Molecular Evolution/MedImmune); Campath 1H/LDP-03 which is a humanized anti CD52 IgG1 antibody (Leukosite); Smart M195 which is a humanized anti-CD33 IgG antibody (Protein Design Lab/Kanebo); RITUXAN™ which is a chimeric anti-CD20 IgG1 antibody (IDEC Pharm/Genentech, Roche/Zettyaku); LYMPHOCIDE™ which is a

humanized anti-CD22 IgG antibody (Immunomedics); ICM3 is a humanized anti-ICAM3 antibody (ICOS Pharm); IDEC-114 is a primate anti-CD80 antibody (IDEC Pharm/Mitsubishi); ZEVALIN™ is a radiolabelled murine anti-CD20 antibody (IDEC/Schering AG); IDEC-131 is a humanized anti-CD40L antibody (IDEC/Eisai);

5   IDEC-151 is a primatized anti-CD4 antibody (IDEC); IDEC-152 is a primatized anti-CD23 antibody (IDEC/Seikagaku); SMART anti-CD3 is a humanized anti-CD3 IgG (Protein Design Lab); 5G1.1 is a humanized anti-complement factor 5 (CS) antibody (Alexion Pharm); D2E7 is a humanized anti-TNF-α antibody (CATIBASF); CDP870 is a humanized anti-TNF-α Fab fragment (Celltech); IDEC-151 is a primatized anti-

10   CD4 IgG1 antibody (IDEC Pharm/SmithKline Beecham); MDX-CD4 is a human anti-CD4 IgG antibody (Medarex/Eisai/Genmab); CDP571 is a humanized anti-TNF-α IgG4 antibody (Celltech); LDP-02 is a humanized anti-α4β7 antibody (LeukoSite/Genentech); OrthoClone OKT4A is a humanized anti-CD4 IgG antibody (Ortho Biotech); ANTOVA™ is a humanized anti-CD40L IgG antibody (Biogen);

15   ANTEGREN™ is a humanized anti-VLA-4 IgG antibody (Elan); and CAT-152 is a human anti-TGF-β$_2$ antibody (Cambridge Ab Tech).

Production of Heterologous Protein by Transgenic Avians

One aspect of the present invention, therefore, concerns transgenic birds, such

20   as chickens, comprising a recombinant nucleic acid molecule and which preferably (though optionally) express a heterologous gene in one or more cells in the animal. Suitable methods for the generation of transgenic avians having heterologous DNA incorporated therein are described, for example, in WO 99/19472 to Ivarie *et al.*; WO 00/11151 to Ivarie *et al.*; and WO 00/56932 to Harvey *et al.*, all of which are

25   incorporated herein by reference in their entirety.

Embodiments of the methods for the production of a heterologous polypeptide by the avian tissue such as the oviduct and the production of eggs which contain heterologous protein involve providing a suitable vector and introducing the vector into embryonic blastodermal cells together with an integrase, preferably phiC31

30   integrase, so that the vector can integrate into the avian genome. A subsequent step involves deriving a mature transgenic avian from the transgenic blastodermal cells

produced in the previous steps. Deriving a mature transgenic avian from the blastodermal cells optionally involves transferring the transgenic blastodermal cells to an embryo and allowing that embryo to develop fully, so that the cells become incorporated into the bird as the embryo is allowed to develop. Another alternative is

5   to transfer a transfected nucleus to an enucleated recipient cell which may then develop into a zygote and ultimately an adult bird. The resulting chick is then grown to maturity.

In an alternative embodiment, the cells of a blastodermal embryo are transfected or transduced with the vector and integrase directly within the embryo. It

10 is contemplated, for example, that the recombinant nucleic acid molecules of the present invention may be introduced into a blastodermal embryo by direct microinjection of the DNA into a stage X or earlier embryo that has been removed from the oviduct. The egg is then returned to the bird for egg white deposition, shell development and laying. The resulting embryo is allowed to develop and hatch, and

15 the chick allowed to mature.

In one embodiment, a transgenic bird of the present invention is produced by introducing into embryonic cells such as, for instance, isolated avian blastodermal cells, a nucleic acid construct comprising an attB recombination site capable of recombining with a pseudo-attP recombination site found within the nuclear genome

20 of the organism from which the cell was derived, and a nucleic acid fragment of interest, in a manner such that the nucleic acid fragment of interest is stably integrated into the nuclear genome of germ line cells of a mature bird and is inherited in normal Mendelian fashion. It is also within the scope of the invention that the targeted cells for receiving the transgene have been engineered to have a heterologous attP

25 recombination site integrated into the nuclear genome of the cells, thereby increasing the efficiency of recognition and recombination with a heterologous attB site.

In either case, the transgenic bird produced from the transgenic blastodermal cells is known as a "founder" Some founders can be chimeric or mosaic birds if, for example, microinjection does not deliver nucleic acid molecules to all of the

30 blastodermal cells of an embryo. Some founders will carry the transgene in the tubular gland cells in the magnum of their oviducts and will express the heterologous

protein encoded by the transgene in their oviducts. If the heterologous protein contains the appropriate signal sequences, it will be secreted into the lumen of the oviduct and onto the yolk of an egg.

Some founders are germ-line founders. A germ-line founder is a founder that carries the transgene in genetic material of its germ-line tissue, and may also carry the transgene in oviduct magnum tubular gland cells that express the heterologous protein. Therefore, in accordance with the invention, the transgenic bird will have tubular gland cells expressing the heterologous protein and the offspring of the transgenic bird will also have oviduct magnum tubular gland cells that express the selected heterologous protein. (Alternatively, the offspring express a phenotype determined by expression of the exogenous gene in a specific tissue of the avian.)

The invention can be used to express, in large yields and at low cost, a wide range of desired proteins including those used as human and animal pharmaceuticals, diagnostics, and livestock feed additives. Proteins such as growth hormones, cytokines, structural proteins and enzymes including human growth hormone, interferon, lysozyme, and β-casein are examples of proteins which are desirably expressed in the oviduct and deposited in eggs according to the invention. Other possible proteins to be produced include, but are not limited to, albumin, α-1 antitrypsin, antithrombin III, collagen, factors VIII, IX, X (and the like), fibrinogen, hyaluronic acid, insulin, lactoferrin, protein C, erythropoietin (EPO), granulocyte colony-stimulating factor (G-CSF), granulocyte macrophage colony-stimulating factor (GM-CSF), tissue-type plasminogen activator (tPA), feed additive enzymes, somatotropin, and chymotrypsin. Immunoglobulins (shown, for example in Example 10 below) and genetically engineered antibodies, including immunotoxins which bind to surface antigens on human tumor cells and destroy them, can also be expressed for use as pharmaceuticals or diagnostics.

In various embodiments of the transgenic bird of the present invention, the expression of the transgene may be restricted to specific subsets of cells, tissues or developmental stages utilizing, for example, *trans*-acting factors acting on the transcriptional regulatory region operably linked to the polypeptide-encoding region of interest of the present invention and which control gene expression in the desired

pattern. Tissue-specific regulatory sequences and conditional regulatory sequences can be used to control expression of the transgene in certain spatial patterns. Moreover, temporal patterns of expression can be provided by, for example, conditional recombination systems or prokaryotic transcriptional regulatory

5   sequences.

The stably modified oviduct cells will express the heterologous polynucleotide and deposit the resulting polypeptide into the egg white of a laid egg. For this purpose, the expression vector will further comprise an oviduct-specific promoter such as ovalbumin or ovomucoid operably linked to the desired heterologous

10   polynucleotide.

Another aspect of the present invention provides a method for the production in an avian of an heterologous protein capable of forming an antibody suitable for selectively binding an antigen. This method comprises a step of producing a transgenic avian incorporating at least one transgene,the transgene encoding at least

15   one heterologous polypeptide selected from an immunoglobulin heavy chain variable region, an immunoglobulin heavy chain comprising a variable region and a constant region, an immunoglobulin light chain variable region, an immunoglobulin light chain comprising a variable region and a constant region, and a single-chain antibody comprising two peptide-linked immunoglobulin variable regions.

20   In one embodiment of this method, the isolated heterologous protein is an antibody capable of selectively binding to an antigen and which may be generated by combining at least one immunoglobulin heavy chain variable region and at least one immunoglobulin light chain variable region, preferably cross-linked by at least one disulfide bridge. The combination of the two variable regions generates a binding site

25   that binds an antigen using methods for antibody reconstitution that are well known in the art.

The present invention also encompasses immunoglobulin heavy and light chains, or variants or derivatives thereof, to be expressed in separate transgenic avians, and thereafter isolated from separate media including serum or eggs, each

30   isolate comprising one or more distinct species of immunoglobulin polypeptide. The method may further comprise the step of combining a plurality of isolated

heterologous immunoglobulin polypeptides, thereby producing an antibody capable of selectively binding to an antigen. In this embodiment, for instance, two or more individual transgenic avians may be generated wherein one transgenic produces serum or eggs having an immunoglobulin heavy chain variable region, or a polypeptide

5   comprising such, expressed therein. A second transgenic animal, having a second transgene, produces serum or eggs having an immunoglobulin light chain variable region, or a polypeptide comprising such, expressed therein. The polypeptides from two or more transgenic animals may be isolated from their respective sera and eggs and combined in vitro to generate a binding site capable of binding an antigen.

10

The present invention is further illustrated by the following examples, which are provided by way of illustration and should not be construed as limiting. The contents of all references, published patents and patents cited throughout the present application are hereby incorporated by reference in their entireties.

15   It will be apparent to those skilled in the art that various modifications, combinations, additions, deletions and variations can be made in the present invention without departing from the scope or spirit of the invention. For instance, features illustrated or described as part of one embodiment can be used in another embodiment to yield a still further embodiment. It is intended that the present invention covers

20   such modifications, combinations, additions, deletions and variations as come within the scope of the appended claims and their equivalents.

Example 1: Phage phiC31 integrase functions in avian cells.

(a) A luciferase vector bearing either an attB (SEQ ID NO: 2 shown in Fig. 10) or attP

25   (SEQ ID NO: 3 shown in Fig. 11) site was co-transfected with an integrase expression vector CMV-C31int (SEQ ID NO: 1) into DF-1 cells, a chicken fibroblast cell line. The cells were passaged several times and the luciferase levels were assayed at each passage.

Cells were passaged every 3-4 days and one third of the cells were harvested

30   and assayed for luciferase. The expression of luciferase was plotted as a percentage of

the expression measured 4 days after transfection. A luciferase expression vector bearing an attP site as a control was also included.

As can be seen in Fig. 2, in the absence of integrase, luciferase expression from a vector bearing attP or attB decreased to very low levels after several days. However, luciferase levels were persistent when the luciferase vector bearing attB was co-transfected with the integrase expression vector, indicating that the luciferase vector had stably integrated into the avian genome.

(b) A drug-resistance colony formation assay was used to quantitate integration efficiency. The puromycin resistance expression vector pCMV-pur was outfitted with an attB (SEQ ID NO: 4 shown in Fig. 12) or an attP (SEQ ID NO: 5 shown in Fig. 13) sites. Puromycin resistance vectors bearing attB sites were cotransfected with phiC31 integrase or a control vector into DF-1 cells. One day after transfection, puromycin was added. Puromycin resistant colonies were counted 12 days post-transfection.

In the absence of co-transfected integrase expression, few DF-1 cell colonies were observed after survival selection. When integrase was co-expressed, multiple DF-1 cell colonies were observed, as shown in Fig. 3. Similar to the luciferase expression experiment, the attB sequence (but not the attP sequence) was able to facilitate integration of the plasmid into the genome. Fig. 3 also shows that phiC31 integrase functions at both 37° Celsius and 41° Celsius. Integrase also functions in quail cells using the puromycin resistance assay, as shown in Fig. 4.

(c) The CMV-pur-attB vector (SEQ ID NO: 4) was also contransfected with an enhanced green fluorescent protein (EGFP) expression vector bearing an attB site (SEQ ID NO: 6 shown in Fig. 14) into DF-1 cells and the phiC31 integrase expression vector CMV-C31int (SEQ ID NO: 1). After puromycin selection for 12 days, the colonies were viewed with UV light to determine the percentage of cells that expressed EGFP. Approximately 20% of puromycin resistant colonies expressed EGFP in all of the cells of the colony, as shown in Fig. 5, indicating that the integrase can mediate multiple integrations per cell.

(d) PhiC31 integrase promoted the integration of large transgenes into avian cells. A puromycin expression cassette comprising a CMV promoter, puromycin resistance gene, polyadenylation sequence and the attB sequence was inserted into a

vector containing a 12.0 kb lysozyme promoter and the human interferon α2b gene (SEQ ID NO: 7 shown in Fig. 15) and into a vector containing a 10.0 kb ovomucoid promoter and the human interferon α2b gene (SEQ ID NO: 8) as shown in Fig. 16.

5   DF-1 cells were transfected with donor plasmids of varying lengths bearing a puromycin resistance gene and an attB sequence in the absence or presence of an integrase expression plasmid. Puromycin was added to the culture media to kill those cells which did not contain a stably integrated copy of the puromycin resistance gene. Cells with an integrated gene formed colonies in the presence of puromycin in 7-12 days. The colonies were visualized by staining with methylene blue and the entire 60

10   mm culture dish was imaged.

PhiC31 integrase mediated the efficient integration of both vectors as shown in Fig. 7.

Example 2: Cell culture methods.

15   DF-1 cells were cultured in DMEM with high glucose, 10% fetal bovine serum, 2 mM L-glutamine, 100 units/ml penicillin and 100 μg/ml streptomycin at 37° Celsius and 5% CO_2. A separate population of DF-1 cells was grown at 41° Celsius. These cells were adapted to the higher temperature for one week before they were used for experiments.

20   Quail QT6 cells were cultured in F10 medium (Gibco) with 5% newborn calf serum, 1% chicken serum heat inactivated (at 55° Celsius for 45 mins), 10 units/ ml penicillin and 10 μg/ml streptomycin at 37° Celsius and 5% CO_2.

Example 3: Selection and Assay Methods

25   (a) *Puromycin selection assay:* About 0.8 x 10^6 DF-1 (chicken) or QT6 (quail) cells were plated in 60 mm dishes. The next day, the cells were transfected as follows:

10 to 50 ng of a donor plasmid and 1 to 10 μg of an Integrase-expressing plasmid DNA were mixed with 150 μl of OptiMEM. 15 μl of DMRIE-C was mixed with 150 μl of OptiMEM in a separate tube, and the mixtures combined and incubated

30   for 15 mins. at room temperature.

While the liposome/DNA complexes were forming, the cells were washed with OptiMEM and 2.5 ml of OptiMEM was added. After 15 minutes, 300 µl of the DNA-lipid mixture was added dropwise to the 2.5 ml of OptiMEM covering the cell layers. The cells were incubated for 4-5 hours at either 37° Celsius or 41° Celsius, 5%

5   CO_2. The transfection mix was replaced with 3 mls of culture media. The next day, puromycin was added to the media at a final concentration of 1 ug/ml, and the media replaced every 2 to 4 days. Puromycin resistant colonies were counted or imaged 10-12 days after the addition of puromycin.

(b) Luciferase assay: Chicken DF-1 or quail QT6 cells (0.8 x 10^6) were plated in 60

10   mm dishes. Cells were transfected as described above. The cells from a plate were transferred to a new 100 mm plate when the plate became confluent, typically on day 3-4, and re-passaged every 3-4 days.

At each time point, one-third of the cells from a plate were replated, and one-third were harvested for the luciferase assay. The cells were pelleted in an eppendorf

15   tube and frozen at –70°C.

The cell pellet was lysed in 200 µl of lysis buffer (25 mM Tris-acetate, pH7.8, 2mM EDTA, 0.5% Triton X-100, 5% glycerol). Sample (5µl) was assayed using the Promega BrightGlo reagent system.

(c) Visualization of EGFP: EGFP expression was visualized with an inverted

20   microscope with FITC illumination [Olympus IX70, 100 W mercury lamp, HQ-FITC Band Pass Emission filter cube, exciter 480/40 nm, emission 535/50 nm, 20X phase contrast objective (total magnification was 2.5 x 10 x 20)].

(d) Staining of cell colonies: After colonies had formed, typically after 7-12 days of culture in puromycin medium, the cells were fixed in 2% formaldehyde, 0.2%

25   glutaraldehyde for 15 mins, and stained in 0.2% methylene blue for 30 mins. followed by several washes with water. The plates were imaged using a standard CCD camera in visible light.

Example 4: Generation of genetically transferred avian cells.

30   Avian stage X blastodermal cells are used as the cellular vector for the transgenes. Stage X embryos are collected and the cells dispersed and mixed with

- 48 -

plasmid DNA.   The transgenes are then introduced to blastodermal cells via electroporation. The cells are immediately injected back into recipient embryos.

The cells are not cultured for any time period to ensure that they remain capable of contributing to the germline of resulting chimeric embryos. However,
5   because there is no culture step, cells that bear the transgene cannot be identified. Typically, only a small percentage of cells introduced to an embryo will bear a stably integrated transgene (0.01 to 1%). To increase the percentage of cells bearing a transgene, therefore, the transgene vector bears an attB site and is co-electroporated with a vector bearing the CMV promoter driving expression of the phiC31 transgene
10   (CMV-C31int (SEQ ID NO: 1)).   The integrase then drives integration of the transgene vector into the nuclear genome of the avian cell and increases the percentage of cells bearing a stable transgene.

(a) Preparation of avian stage X blastodermal cells:

i)   Collect fertilized eggs from Barred Rock or White leghorn chickens
15   (*Gallus gallus*) or quail (*Japonica coturnix*) within 48 hrs. of laying;

ii)   Use 70% ethanol to clean the shells;

iii) Crack the shells and open the eggs;

iv)   Remove egg whites by transferring yolks to opposite halves of shells, repeating to remove most of the egg whites;

20   v)   Put egg yolks with embryo discs facing up into a 10cm petri dish;

vi)   Use an absorbent tissue to gently remove egg white from the embryo discs;

vii) Place a Whatman filter paper 1 ring over the embryos;

viii) Use scissors to cut the membranes along the outside edge of the paper ring while gently lifting the ring/embryos with a pair of tweezers;

25   ix)   Insert the paper ring with the embryos at a 45 degrees angle into a petri dish containing PBS-G solution at room temperature;

x)   After ten embryo discs are collected, gently wash the yolks from the blastoderm discs using a Pasteur pipette under a stereo microscope;

xi)   Cut the discs by a hair ring cutter (a short piece of human hair is bent into a
30   small loop and fastened to the narrow end of a Pasteur pipette with Parafilm);

xii)   Transfer the discs to a 15 ml sterile centrifuge tube on ice;

xiii)   Place 10 to 15 embryos per tube and allow to settle to the bottom (about 5 mins.);

xiv)   Aspirate the supernatant from the tube;

xv)   Add 5 mls of ice-cold PBS without Ca^{++} and Mg^{++}, and gently pipette 4 to 5 times using a 5 mls pipette;

xvi)   Incubate in ice for 5-7 mins. to allow the blastoderms to settle, and aspirate the supernatant;

xvii)   Add 3 mls of ice cold 0.05% trypsin/0.02% ETDA to each tube and gently pipette 3 to 5 times using a 5 ml pipette;

xviii)   Put the tube in ice for 5 mins. and then flick the tube by finger 40 times. Repeat;

xix)   Add 0.5 mls FBS and 3-5 mls BDC medium to each tube and gently pipette 5-7 times using a 5 ml pipette;

xx)   Spin at 500 rpm (RCF 57 x g) at 4° Celsius for 5 mins;

xxi)   Remove the supernatant and add 2 mls ice cold BDC medium into each tube; and

xxii)   Resuspend the cells by gently pipetting 20-25 times; and

xxiii)   Determine the cell titer by hemacytometer and ensure that about 95% of all BDCs are single cells, and not clumped.

(b) Transfection of linearized plasmids into blastodermal cells by small scale electroporation:

i)   Centrifuge the blastodermal cell suspension from step (xxiii) above at RCF 57 x g, 4° Celsius, for 5 mins;

ii)   Resuspend cells to a density of 1-3 x 10^6 per ml with PBS without Ca^{2+} and Mg^{2+};

iii)   Add linearized DNA, 1-30 µg per 1-3 x 10^5 blastodermal cells in an eppendorf tube at room temperature. Add equimolar molar amounts of the non-linearized transgene plasmid bearing an attB site, and an integrase expression plasmid;

iv)   Incubate at room temperature for 10 mins;

v)   Aliquot 100 µl of the DNA-cell mixture to a 0.1 cm cuvette at room temperature;

vi)   Electroporate at 240 V and 25 µFD (or 100 V and 125 µFD for quail cells) using, for example, a Gene Pulser II™ (BIO-RAD).

5   vii)   Incubate the cuvette at room temperature for 1-10 mins.

viii)   Before the electroporated cells are injected into a recipient embryo, they are transferred to a eppendorf tube at room temperature. The cuvette is washed with 350 µl of media, which is transferred to the eppendorf, spun at room temperature and re-suspended in 0.01-0.3 ml medium;

10   ix)   Inject 1-10 µl of cell suspension into the subgerminal cavity of an non-irradiated or, preferably, an irradiated (e.g., with 300-900 rads) stage X egg. Shell and shell membrane are removed and, after injection, resealed according to U.S. Patent No. 6,397,777 incorporated herein by reference in its entirety; and

15   x)   The egg is then incubated to hatching.

(c) Blastodermal Cell Culture Medium:

i)   409.5 mls DMEM with high glucose, L-glutamine, sodium pyruvate, pyridoxine hydrochloride;

ii)   5 mls Men non-essential amino acids solution, 10 mM;

20   iii)   5 mls Penicillin-streptomycin 5000 U/ml each;

iv)   5 mls L-glutamine, 200 mM;

v)   75 mls fetal bovine serum; and

vi)   0.5 mls β-mercaptoethanol, 11.2mM.

25   **Example 5: Transfection of stage X embryos with attB plasmids**

(a) DNA-PEI: Twenty-five µg of a phage phiC31 integrase expression plasmid (pCMV-int), and 25 µg of a luciferase-expressing plasmid (pβ-actin-GFP-attB) are combined in 200 µl of 28 mM Hepes (pH 7.4). The DNA/Hepes is mixed with an equal volume of PEI which has been diluted 10-fold with water. The DNA/Hepes/PEI

30   is incubated at room temperature for 15 mins Three to seven µl of the complex are injected into the subgerminal cavity of windowed stage X white leghorn eggs which

are then sealed and incubated as described in U.S. Patents No. 6,397,777. The complexes will also be incubated with blastodermal cells isolated from stage X embryos which are subsequently injected into the subgerminal cavity of windowed irradiated stage X white leghorn eggs. Injected eggs are sealed and incubated as

5   described above.

(b) Adenovirus-PEI:

Two μg of a phage phiC31 integrase expression plasmid (pCMV-int), 2 μg of a GFP expressing plasmid (pβ-actin-GFP-attB) and 2 μg of a luciferase expressing plasmid (pGLB) were incubated with 1.2 μl of JetPEI™ in 50 μl of 20 mM Hepes buffer

10   (pH7.4). After 10 mins at 25°C, 3×10^9 adenovirus particles (Ad5-Null, Qbiogene) were added and the incubation continued for an additional 10 mins. Embryos are transfected *in ovo* or *ex ovo* as described above.

Example 6: Stage I cytoplasmic injection

15   Production of transgenic chickens by cytoplasmic DNA injection using DNA injection directly into the germinal disk as described in Sang *et al., Mol. Reprod. Dev.*, 1: 98-106 (1989); Love *et al., Biotechnology*, 12: 60-63 (1994) incorporated herein by reference in their entireties.

           In the method of the present invention, fertilized ova, and preferably stage I

20   embryos, are isolated from euthanized hens 45 mins. to 4 hrs. after oviposition of the previous egg. Alternatively, eggs were isolated from hens whose oviducts have been fistulated according to the techniques of Gilbert & Wood-Gush, *J. Reprod. Fertil.*, 5: 451-453 (1963) and Pancer *et al., Br. Poult. Sci.*, 30: 953-7 (1989) incorporated herein in their entireties.

25   An isolated ovum was placed in dish with the germinal disk upwards. Ringer's buffer medium was then added to prevent drying of the ovum. Any suitable microinjection assembly and methods for microinjecting and reimplanting avian eggs are useful in the method of cytoplasmic injection of the present invention. A particularly suitable apparatus and method for use in the present invention is described

30   in U.S. Patent Application Serial No: 09/919,143 ("the '143 Application) and incorporated herein by reference in its entirety. The avian microinjection system

described in the '143 Application allowed the loading of a DNA solution into a micropipette, followed by prompt positioning of the germinal disk under the microscope and guided injection of the DNA solution into the germinal disk. Injected embryos could then be surgically transferred to a recipient hen as described, for

5   example, in Olsen & Neher, *J. Exp. Zool.*, 109: 355-66 (1948) and Tanaka *et al.*, *J. Reprod. Fertil.*, 100: 447-449 (1994). The embryo was allowed to proceed through the natural *in vivo* cycle of albumin deposition and hard-shell formation. The transgenic embryo is then laid as a hard-shell egg which was incubated until hatching of the chick. Preferably, injected embryos were surgically transferred to recipient

10   hens via the ovum transfer method of Christmann *et al.* in PCT/US01/26723, the contents of which are incorporated by reference in its entirety, and hard shell eggs were incubated and hatched.

Approximately 25 nl of DNA solution with either integrase mRNA or protein were injected into a germinal disc of stage I White Leghorn embryos obtained 90

15   minutes after oviposition of the preceding egg. Typically the concentration of integrase mRNA used was 100 ng/μl, and the concentration of integrase protein was 66 ng/μl.

To synthesize the integrase mRNA, a plasmid template encoding the integrase protein was linearized at the 3' end of the transcription unit. mRNA was synthesized,

20   capped and a polyadenine tract added using the mMESSAGE mMACHINE T7 Ultra Kit™ (Ambion, Austin, TX). The mRNA was purified by extraction with phenol and chloroform and precipitiated with isopropanol. The integrase protein was expressed in *E. coli* and purified as described by Thorpe *et al.*, *Mol. Microbiol.*, 38: 232-241 (2000).

25   A plasmid encoding for the integrase protein is transfected into the target cells. However, since the early avian embryo transcriptionally silent until it reaches about 22,000 cells, injection of the integrase mRNA or protein was expected to result in better rates of transgenesis, as shown in the Table 1 below.

The chicks produced by this procedure were screened for the presence of the

30   injected transgene using a high throughput PCR-based screening procedure as described in Harvey *et al.*, *Nature Biotech.*, 20: 396-399 (2002).

Table 1: Summary of cytoplasmic injection results using different integrase strategies

Experimental group	*Ovum* transfers	Hard shells produced (%)	Chicks hatched (%) *	Transgenic chicks (%) ‡
No Integrase	5164	3634 (70%)	500 (14%)	58 (11.6%)
Integrase mRNA	1109	833 (75%)	115 (13.8%)	19 (16.5%)
Integrase protein	374	264 (70.6%)	47(17.8%)	16 (34%)

* :   Percentages based on the number of hard shells

‡ :   Percentages based on the number of hatched birds

Example 7: Characterization of phiC31 integrase-mediated integration sites in the chicken genome.

To characterize phiC31-mediated integration into the chicken genome, a plasmid rescue method was used to isolate integrated plasmids from transfected and selected chicken fibroblasts. Plasmid pCR-XL-TOPO-CMV-pur-attB (SEQ ID NO: 10, shown in Fig. 18) does not have *Bam*H I or *Bgl* II restriction sites. Genomic DNA from cells transformed with pCR-XL-TOPO-CMV-pur-attB was cut with *Bam*H I or *Bgl* II (either or both of which would cut in the flanking genomic regions) and religated so that the genomic DNA surrounding the integrated plasmid would be captured into the circularized plasmid. The flanking DNA of a number of plasmids were then sequenced.

DF-1 cells (chicken fibroblasts), 4 x 10^5 were transfected with 50 ng of pCR-XL-TOPO-CMV-pur-attB and 1 μg of pCMV-int. The following day, the culture medium was replaced with fresh media supplemented with 1 μg/ml puromycin. After 10 days of selection, several hundred puromycin-resistant colonies were evident. These were harvested by trypsinzation, pooled, replated on 10 cm plates and grown to confluence. DNA was then extracted.

Isolated DNA was digested with *Bam*H I and *Bgl* II for 2-3 hrs, extracted with phenol:chloroform:isoamyl alcohol chloroform:isoamyl alcohol and ethanol

precipitated. T4 DNA ligase was added and the reaction incubated for 1 hr at room temperature, extracted with phenol:chloroform:isoamyl alcohol and chloroform:isoamyl alcohol, and precipitated with ethanol. 5 µl of the DNA suspended in 10µl of water was electroporated into 25 µl of Genehogs™ (Invitrogen) in an 0.1 cm cuvette using a GenePulser II (Biorad) set at 1.6 kV, 100 ohms, 25 uF and plated on Luria Broth (LB) plates with 5 µg/ml phleomycin (or 25 µg/ml zeocin) and 20 µg/ml kanamycin. Approximately 100 individual colonies were cultured, the plasmids extracted by standard miniprep techniques and digested with *Xba* I to identify clones with unique restriction fragments.

Thirty two plasmids were sequenced with the primer attB-for (5'-TACCGTCGACGATGTAGGTCACGGTC-3') (SEQ ID NO: 12) which allows sequencing across the crossover site of attB and into the flanking genomic sequence. All of plasmids sequenced had novel sequences inserted into the crossover site of attB, indicating that the clones were derived from plasmid that had integrated into the chicken genome via phiC31 integrase-mediated recombination.

The sequences were compared with sequences at GenBank using Basic Local Alignment Search Tool (BLAST). Most of the clones harbored sequences homologous to *Gallus* genomic sequences in the TRACE database.

Example 3: Insertion of a wild-type attP site into the avian genome augments integrase-mediated integration and transgenesis.

The chicken B-cell line DT40 cells (Buerstedde *et al.*, *E.M.B.O. J.*, 9: 921-927 (1990)) are useful for studying DNA integration and recombination processes (Buerstedde & Takeda, *Cell*, 67:179-88 (1991)). DT40 cells were engineered to harbor a wild-type attP site isolated from the *Streptomyces* phage phiC31. Two independent cell lines were created by transfection of a linearized plasmid bearing an attP site linked to a CMV promoter driving the resistance gene to G418 (DT40-NLB-attP) or bearing an attP site linked to a CMV promoter driving the resistance gene for puromycin (DT40-pur-attP). The transfected cells were cultured in the presence of G418 or puromycin to enrich for cells bearing an attP sequence stably integrated into the genome.

A super-coiled luciferase vector bearing an attB (SEQ ID NO: 2 shown in Fig. 10) was co-transfected, together with an integrase expression vector CMV-C31int (SEQ ID NO: 1) or a control, non-integrase expressing vector (CMV-BL) into wild-type DT40 cells and the stably transformed lines DT40-NLB-attP and DT40-pur-attP.

5   Cells were passaged at 5, 7 and 14 days post-transfection and about one third of the cells were harvested and assayed for luciferase. The expression of luciferase was plotted as a percentage of the expression measured 5 days after transfection. As can be seen in Fig. 21, in the absence of integrase, or in the presence of integrase but in the DT40 cells lacking an inserted wild-type attP site, luciferase expression from a

10   vector bearing attB progressively decreased to very low levels. However, luciferase levels were persistent when the luciferase vector bearing attB was co-transfected with the integrase expression vector into the attP bearing cell lines DT40-NLB-attP and DT40-pur-attP. Inclusion of an attP sequence in the avian genome augments the level of integration efficiency beyond that afforded by the utilization of endogenous

15   pseudo-attP sites.

Example 9: Generation of attP transgenic cell line and birds using an NLB vector

The NLB-attP retroviral vector can be injected into stage X chicken embryos

20   laid by pathogen-free hens. A small hole is drilled into the egg shell of a freshly laid egg, the shell membrane cut away and the embryo visualized by eye. With a drawn needle attached to a syringe, 1 to 10 μl of concentrated retrovirus, approximately 2.5 x 10^5 IU, is injected into the subgerminal cavity of the embryo. The egg shell is resealed with a hot glue gun. Suitable methods for the manipulation of avian eggs, including

25   opening and resealing hard shell eggs are described in U.S. Patent Serial Nos: 5,897,998 and 6,397,777 which are herein incorporated by reference in their entireties.

Typically, 25% of embryos hatch 21 days later. The chicks are raised to sexual maturity and semen samples are taken. Birds that have a significant level of the transgene in sperm DNA will be identified, typically by a PCR-based assay. Ten to

30   25% of the hatched roosters will be able to give rise to G1 transgenic offspring, 1 to 20% of which may be transgenic. DNA extracted from the blood of G1 offspring is

analyzed by PCR and Southern analysis to confirm the presence of the intact transgene. Several lines of transgenic roosters, each with a unique site of attP integration, are then bred to non-transgenic hens, giving 50% of G2 transgenic offspring. Transgenic G2 hens and roosters from the same line can be bred to produce

5   G3 offspring homozygous for the transgene. Homozygous offspring will be distinguished from hemizygous offspring by quantitative PCR. The same procedure can be used to integrate an attB or attP site into transgenic birds.

Example 10: Expression of immunoglobulin chain polypeptides by transgenic

10   chickens

Bacterial artificial chromosomes (BACs) containing a 70 kbp segment of the chicken ovomucoid gene with the light and heavy chain cDNAs for a human monoclonal antibody inserted along with an internal ribosome entry site into the 3' untranslated region of the ovomucoid gene were equipped with the attB sequence.

15   The heavy and light chain cDNAs were inserted into separate ovomucoid BACs such that expression of an intact monoclonal antibody requires the presence of both BACs in the nucleus.

Several hens produced by coinjection of the attB-bearing ovomucoid BACs and integrase-encoding mRNA into stage I embryos produced intact monoclonal

20   antibodies in their egg white.. One hen, which had a high level of the light chain ovomucoid BAC in her blood DNA as determined by quantitative PCR particularly expressed the light chain portion of the monoclonal antibody in the egg white at a concentration of 350 nanograms per ml, or approximately 12 µg per egg.

What is Claimed Is:

1.   A method for genetically transforming an avian cell, comprising:

                delivering to an avian cell having a first recombination site a nucleic
        acid molecule comprising a second recombination site;

5   delivering a source of integrase activity to the avian cell; and

                maintaining the avian cell under conditions suitable for the integrase to
        mediate recombination between the first and the second recombination
        sites, thereby integrating the nucleic acid molecule into the nuclear
        genome of the avian cell.

10

2.   The method of Claim 1, wherein the avian cell is selected from the group
        consisting of a stage I cell, a stage X blastodermal cell, a primordial germline
        cell and an oviduct cell.

15 3.   The method of Claim 1, wherein the integrase is a serine recombinase

4.   The method of Claim 1, wherein the integrase is the bacteriophage phiC31
        integrase.

20 5.   The method of Claim 1, wherein the source of integrase activity is delivered to
        the avian cell as a polypeptide.

6.   The method of Claim 1, wherein the source of integrase activity is expressed
        from a polynucleotide molecule delivered to the avian cell.

25

7.   The method of Claim 6, wherein the polynucleotide molecule is an mRNA.

8.   The method of Claim 6, wherein the polynucleotide molecule is an expression
        vector.

30

9.   The method of Claim 8, wherein the expression vector comprises a promoter operably linked to a region encoding the integrase, wherein the promoter is selected from a CMV promoter and an RSV promoter.

5   10.   The method of Claim 9, wherein the expression vector comprises a nucleotide sequence selected from SEQ ID NOs: 1 and 9.

11.   The method of Claim 1, wherein the nucleic acid molecule comprises an isolated gene expression control region operably linked to a region encoding a

10   polypeptide selected to be expressed by an avian cell.

12.   The method of Claim 1, wherein the nucleic acid molecule is an expression vector.

15   13.   The method of Claim 12, wherein the expression vector is selected from the group consisting of a plasmid, an artificial chromosome, an isolated eukaryotic chromosome, an adenovirus vector and a lentivirus vector.

14.   The method of Claim 11, wherein the gene expression control region is

20   isolated from an avian gene selected from the group consisting of an ovalbumin gene, a lysozyme gene and an ovomucoid gene.

15.   The method of Claim 11, wherein the nucleic acid molecule further comprises an IRES operably linked to a region encoding a second polypeptide selected to

25   be expressed by an avian cell.

16.   The method of Claim 11, wherein the region encoding the polypeptide is codon-optimized for expression of the polypeptide by the avian cell.

30   17.   The method of Claim 15 wherein the region encoding the second polypeptide is codon-optimized for expression of the polypeptide by the avian cell.

18.   The method of Claim 1, wherein the first recombination site is a attP site having at least 25% identity to the sequence according to SEQ ID NO: 11.

5   19.   The method of Claim 1, wherein the first recombination site is a attP site having at least 75% identity to the sequence according to SEQ ID NO: 11.

20.   The method of Claim 1, wherein the first recombination site is a attP site according to SEQ ID NO: 11.

10

21.   The method of Claim 1, wherein the first recombination site is a pseudo-attP site.

22.   The method of Claim 1, wherein the nucleic acid molecule and the source of
15   integrase activity are sequentially delivered to the isolated avian cell.

23.   The method of Claim 1, wherein the nucleic acid molecule and the source of integrase activity are concurrently delivered to the isolated avian cell.

20   24.   The method of Claim 1, wherein the avian cell is a chicken cell.

25.   The method of Claim 1, wherein the avian cell is a quail cell.

26.   The method of Claim 1, wherein the nucleic acid molecule and the source of
25   integrase activity are delivered to the avian cell by a method or combination of methods selected from injection, electroporation, DNA-PEI, virus-PEI, microcell fusion, artificial viral coats, polylysine conjugation and lipofection.

27.   The method of Claim 1, wherein the avian cell is an oviduct cell in the oviduct
30   of an avian and further comprising the steps of:

before delivering the nucleic acid molecule and source of integrase activity, surgically exposing the luminal surface of the oviduct of an avian;

delivering the nucleic acid molecule and the source of integrase activity to the oviduct cell by depositing the nucleic acid molecule and the source of integrase activity on the luminal surface;

applying an electroporation pulse across the wall of the oviduct, wherein the electroporation pulse is directed to deliver the nucleic acid molecule and the source of integrase activity to recipient oviduct cells, whereby the integrase mediates integration of the nucleic acid molecule into the genome of a recipient cell;

surgically closing the oviduct; and

maintaining the avian under conditions suitable for the expression by the recipient oviduct cells of a polypeptide encoded by the nucleic acid molecule and deposition of said polypeptide into the white of a laid hard shell egg.

28. A genetically transformed avian cell, and progeny thereof, generated by a of the methods according to Claims 1.

29. A genetically transformed avian, and progeny thereof, generated by the method according to Claim 27.

30. A method for generating a genetically transformed avian, comprising:

genetically transforming an avian cell by a method according to Claim 1;

delivering the genetically transformed avian cell to an early stage avian embryo, thereby generating a genetically transformed embryo; and

maintaining the avian embryo under conditions suitable for the embryo to develop and hatch as a transgenic chick.

31.   The method of Claim 26, wherein the genetically transformed embryo is delivered to a recipient female avian and laid as a hard shell egg.

5   32.   A method of producing a polypeptide comprising expressing the polypeptide in a transgenic avian cell, said cell comprising a heterologous nucleic acid molecule that expresses the polypeptide, wherein the heterologous nucleic acid molecule is integrated into the nuclear genome and flanked by halves of an integration site recognized by an integrase.

10

33.   A polypeptide of interest expressed by a genetically transformed avian cell, and progeny thereof, the cell comprising a heterologous nucleic acid molecule integrated into the nuclear genome thereof by an integrase, and wherein the nucleic acid molecule comprises a region encoding the polypeptide of interest

15   and an operably linked promoter region.

34.   The polypeptide according to Claim 33, wherein the polypeptide is an immunoglobulin chain.

20   35.   The polypeptide according to Claim 33, wherein the polypeptide is a cytokine.

36.   The polypeptide according to Claim 33, wherein the polypeptide is deposited in the white of an avian egg.

25   37.   The polypeptide according to Claim 33, wherein the polypeptide is expressed by a transgenic bird comprising the genetically transformed avian cell.

38.   The polypeptide according to Claim 37, wherein the genetically transformed avian cell, and the progeny thereof, are generated according to the methods of

30   Claim 1.

39. The polypeptide according to Claim 37, wherein the transgenic bird is generated according to the method of Claim 30.

40. An avian egg white comprising a heterologous polypeptide produced according to the method according to the Claim 32.

41. An avian egg comprising a heterologous polypeptide produced according to the method according to the Claim 32.

42. A method for generating a genetically modified avian cell, and progeny thereof, using a tagged chromosome, said method comprising the steps of:

    providing an isolated modified chromosome comprising a lac operator region and a first recombination site;

    delivering the modified chromosome to a avian cell, thereby generating a trisomic avian cell;

    delivering to the avian cell a source of a tagged polypeptide comprising a fluorescent domain and a lac repressor domain;

    delivering a source of integrase activity to the avian cell;

    delivering a polynucleotide comprising a second recombination site and a region encoding a polypeptide to the avian cell;

    maintaining the avian cell under conditions suitable for the integrase to mediate recombination between the first and second recombination sites, thereby integrating the polynucleotide into the modified chromosome and generating a genetically modified avian cell;

    expressing the tag polypeptide by the avian cell;

    allowing the tag polypeptide to bind to the modified chromosome so as to label the modified chromosome; and

    isolating the modified chromosome by selecting modified chromosomes having a tag polypeptide bound thereto.

43.   The method according to Claim 42, further comprising the step of delivering the modified chromosome to a second avian cell to generate a trisomic avian cell.

44.   A modified avian chromosome comprising a lac operator region and a recombination site.

45.   The avian chromosome according to Claim 44, wherein the lac operator region is a concatamer of lac operators.

46.   The avian chromosome according to Claim 44, wherein the recombination site is selected from an att B or an att P site.

47.   The avian chromosome according to Claim 44, further comprising a region encoding a tag polypeptide and an operably linked promoter.

48.   The avian chromosome according to Claim 47, wherein the tag polypeptide comprises a fluorescent domain and a lac repressor domain.

49.   The avian chromosome according to Claim 44, wherein the florescent domain is GFP.

50.   The avian chromosome according to Claim 45, further comprising a heterologous polynucleotide integrated into the first recombination site.

51.   The avian chromosome according to Claim 50, wherein the heterologous polynucleotide comprises a region encoding a polypeptide operably linked to a promoter.

52.   The avian chromosome according to Claim 50, wherein the promoter is an avian promoter selected from the gene expression control region of an onomucoid gene, a lysozyne gene and an ovalbumin gene.
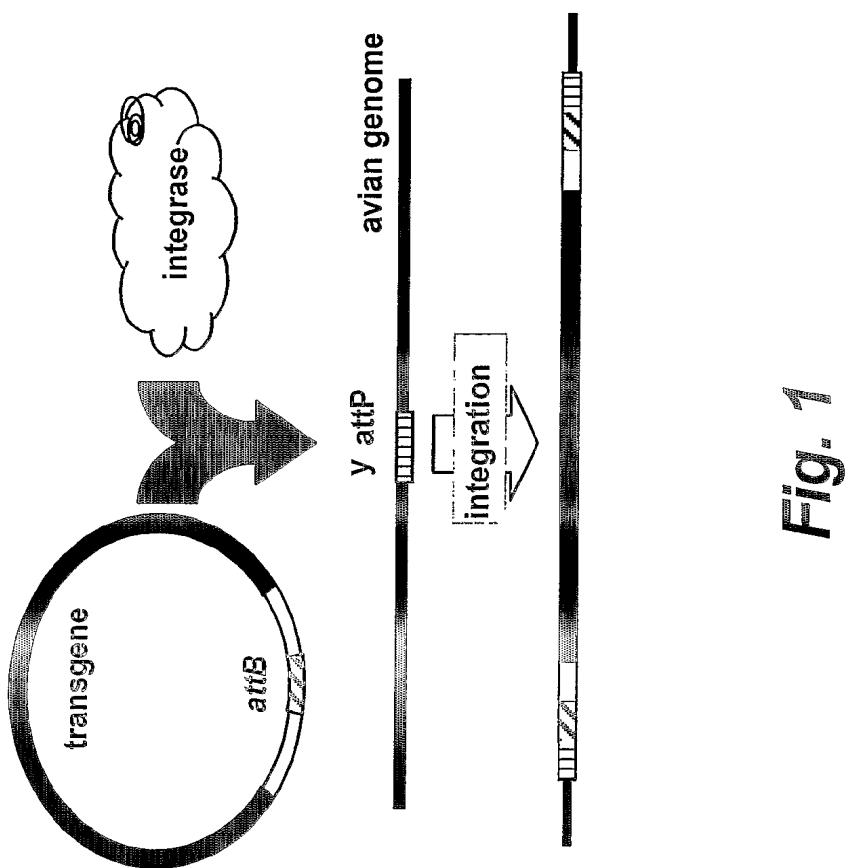
Fig. 1
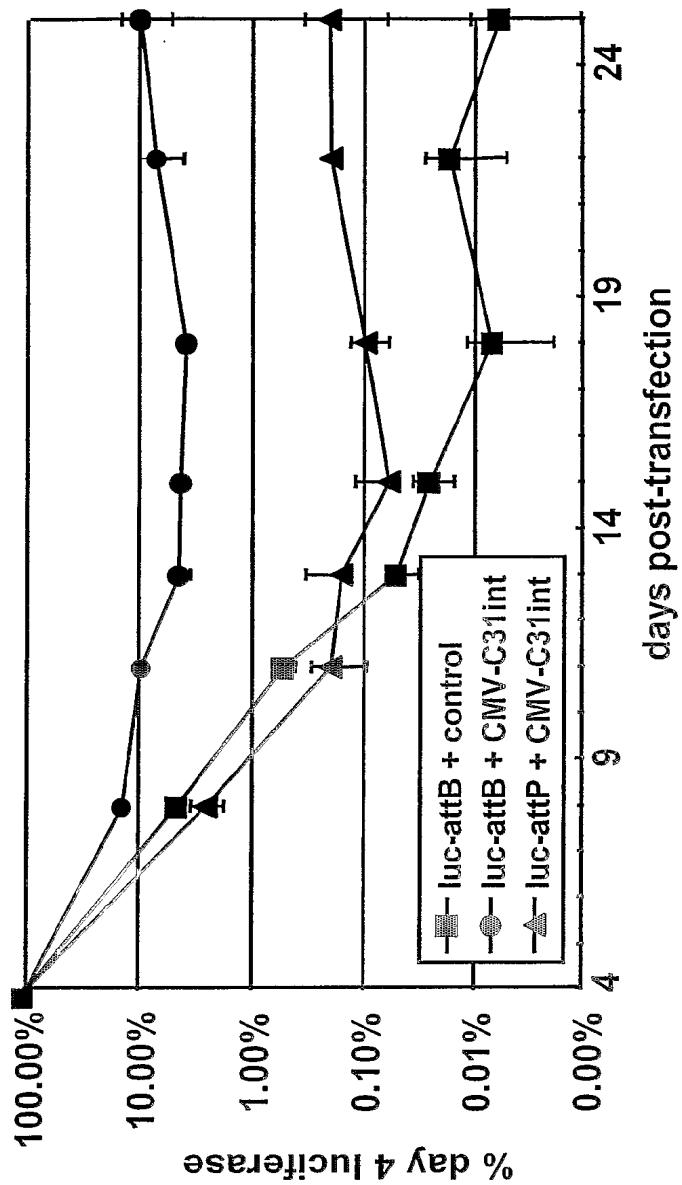
Fig. 2
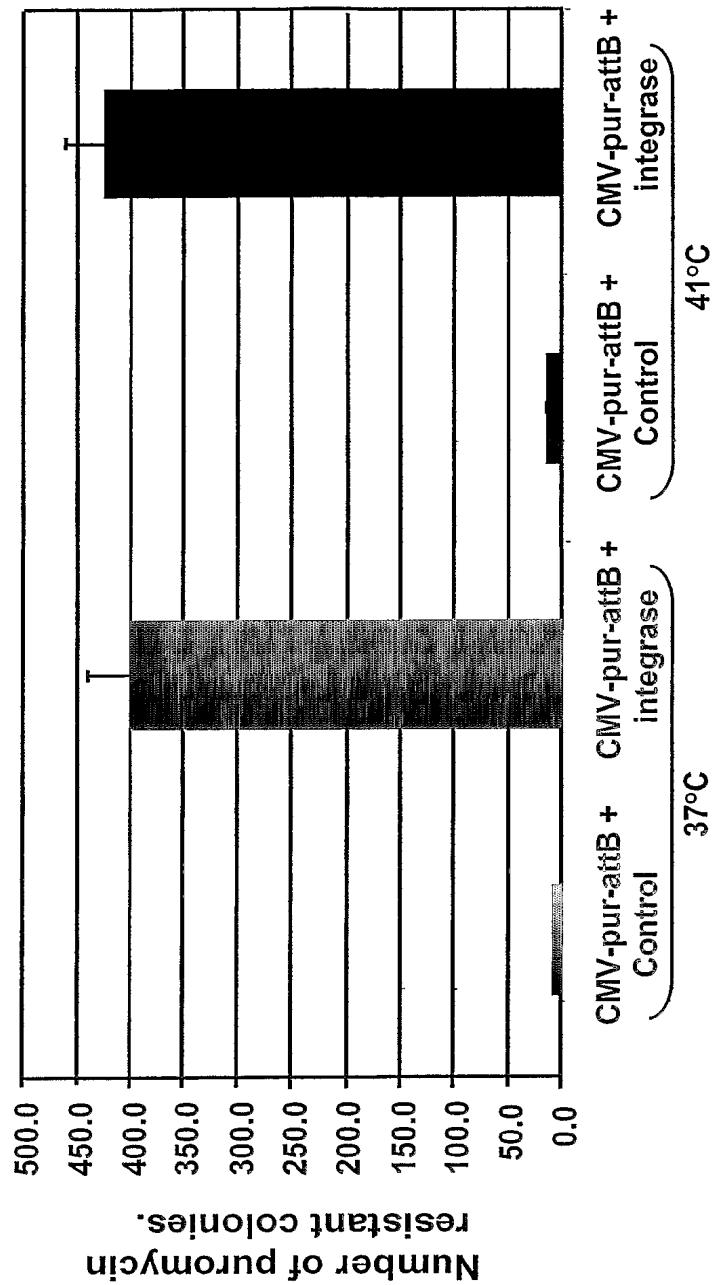
Fig. 3
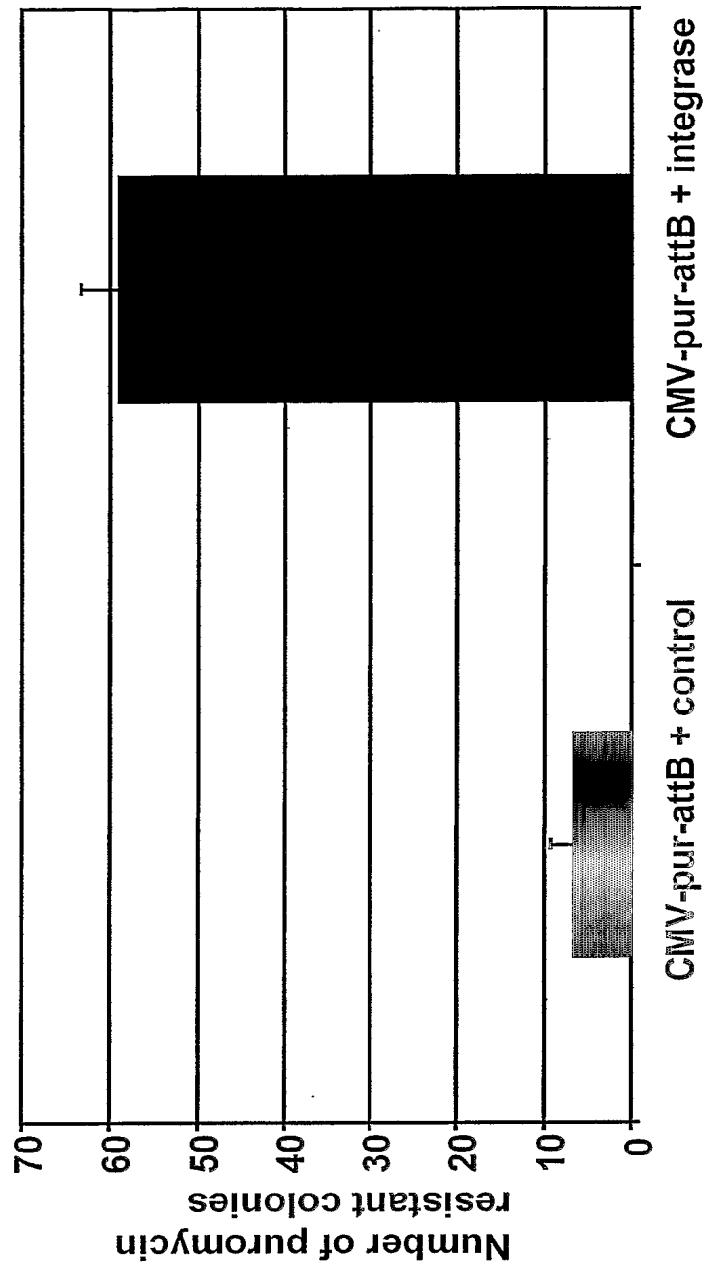
Fig. 4
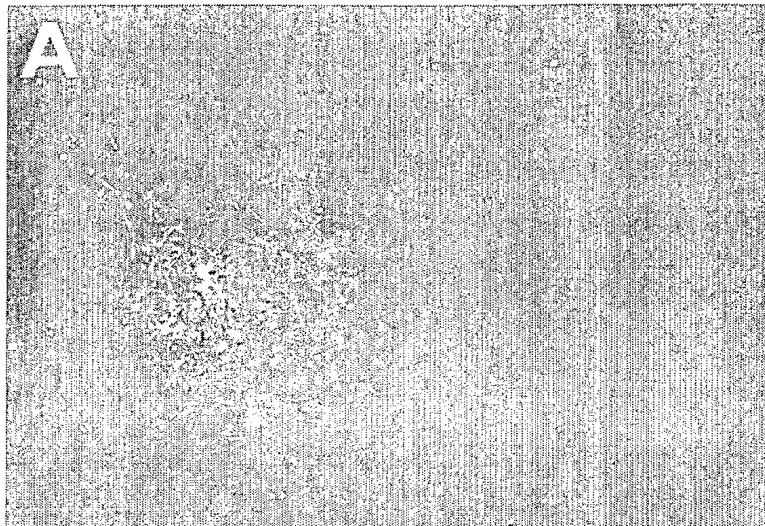
Fig. 5
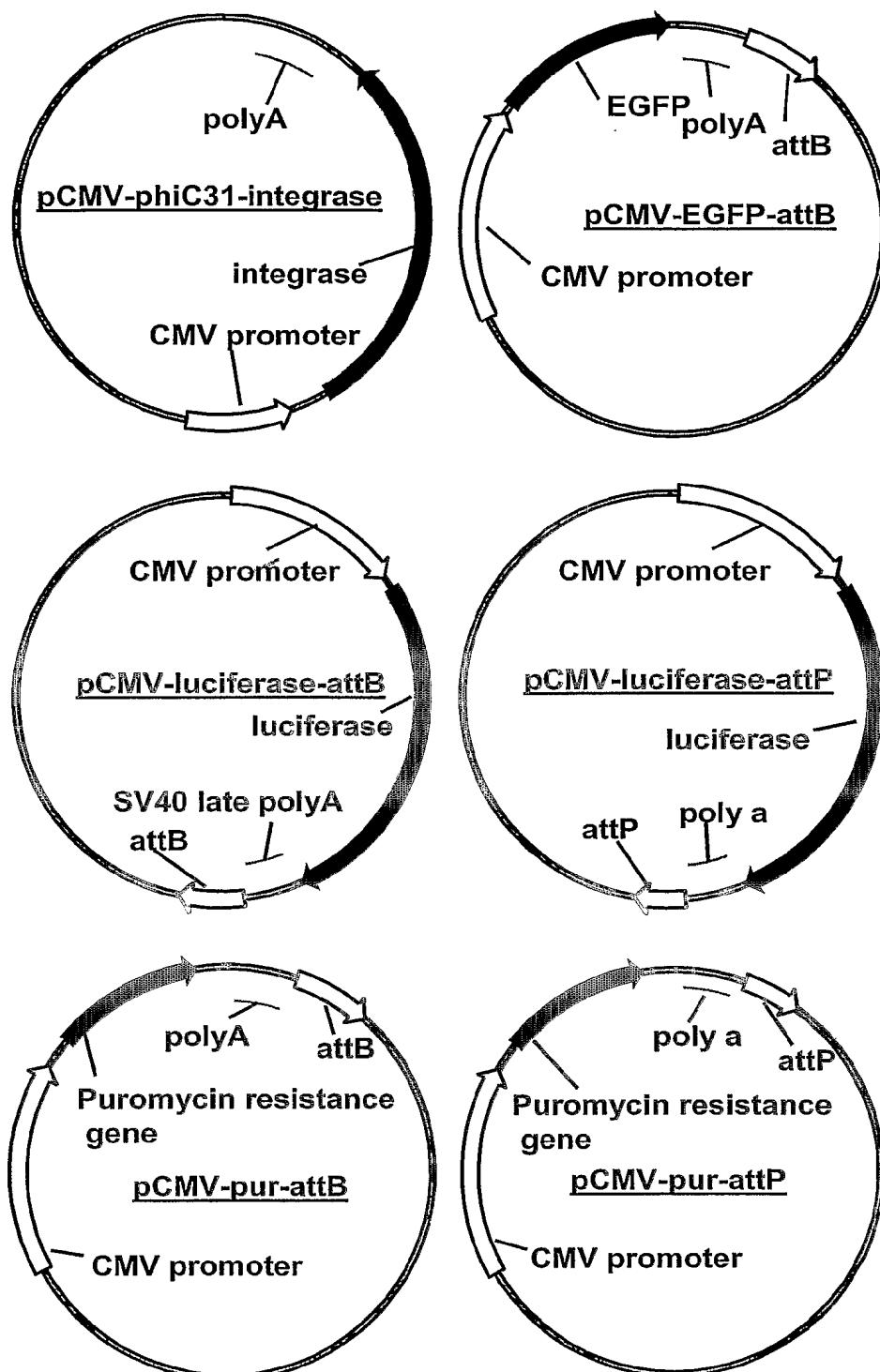
Fig. 6
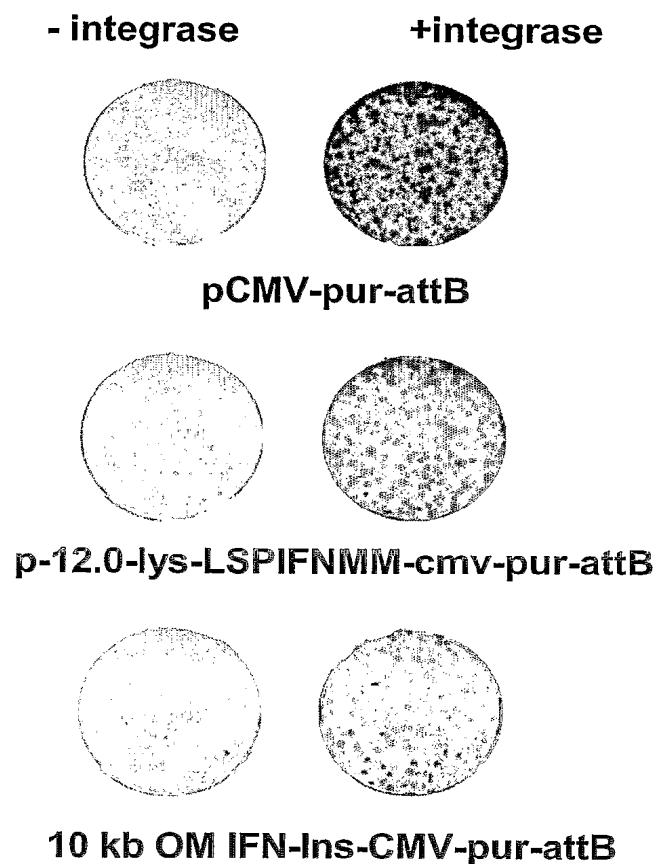
Fig. 7
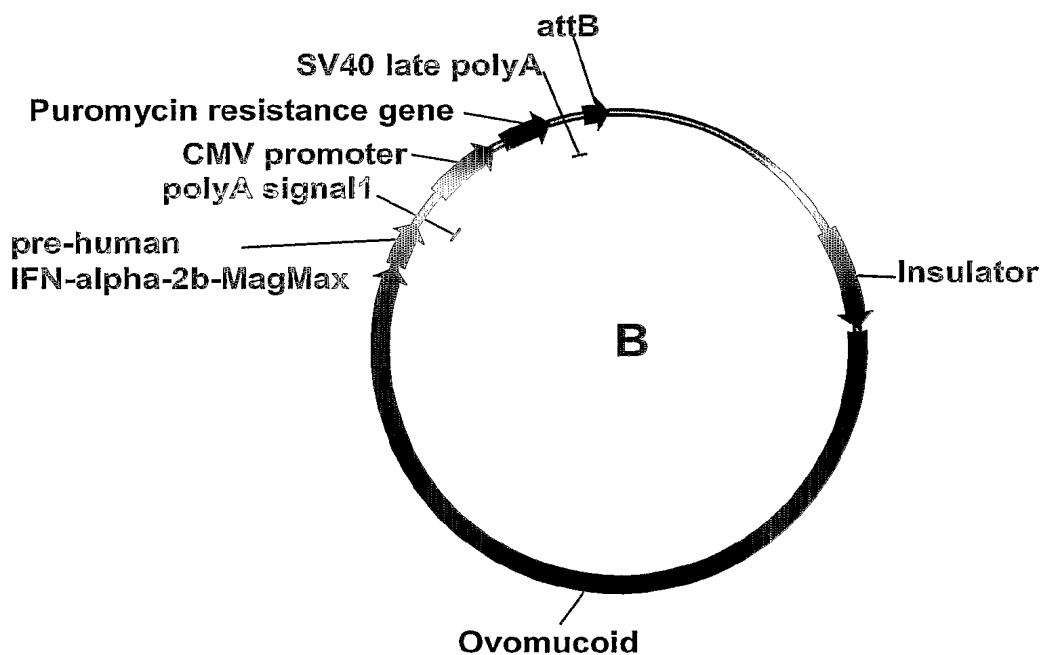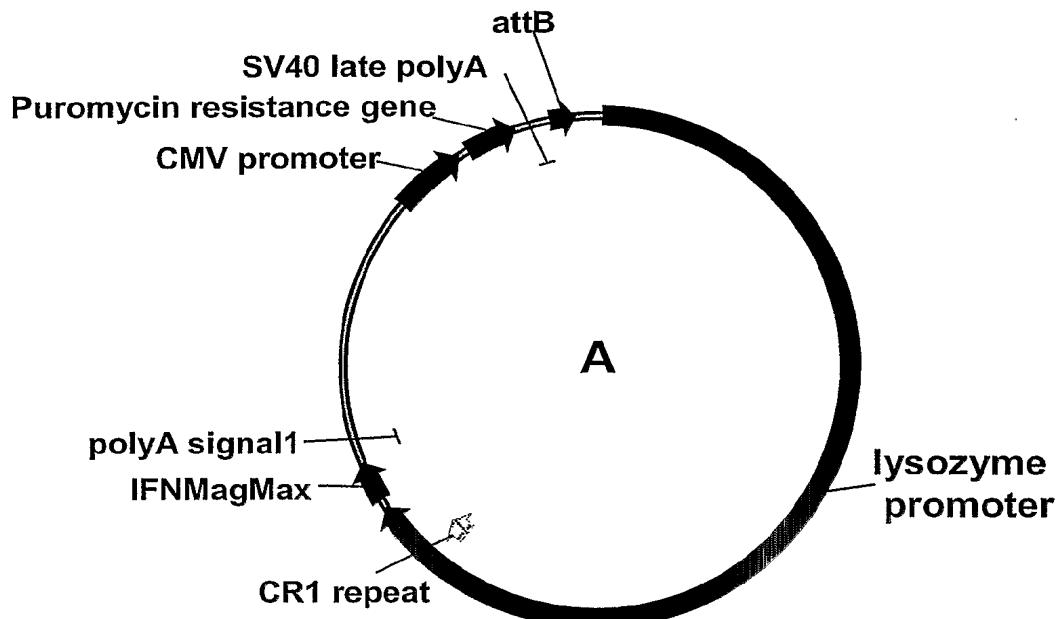
Fig. 8

pCMV-C31int (SEQ ID NO: 1)

```
CATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATT
ACGCCAGCCAATACGCAAACCGCCTCTCCCGCGCGTTGGCCGATTCATTAATGCAGGATCG
ATCCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAA
AAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAA
TAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGG
AGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTATGGCTGATTATGATCATGAACAG
ACTGTGAGGACTGAGGGGCCTGAAATGAGCCTTGGGACTGTGAATCTAAAATACACAAACAA
TTAGAATCACTAGCTCCTGTGTATAATATTTTCATAAATCATACTCAGTAAGCAAAACTCTC
AAGCAGCAAGCATATGCAGCTAGTTTAACACATTATACACTTAAAAATTTTATATTTACCTT
AGAGCTTTAAATCTCTGTAGGTAGTTTGTCCAATTATGTCACACCACAGAAGTAAGGTTCCT
TCACAAAGATCCCAAGCTAGCTTATAATACGACTCACTATAGGGAGAGAGCTATGACGTCGC
ATGCACGCGTAAGCTTGGGCCCCTCGAGGGATCCGGGTGTCTCGCTACGCCGCTACGTCTTC
CGTGCCGTCCTGGGCGTCGTCTTCGTCGTCGTCGGTCGGCGGCTTCGCCCACGTGATCGAAG
CGCGCTTCTCGATGGGCGTTCCCTGCCCCCTGCCCGTAGTCGACTTCGTGACAACGATCTTG
TCTACGAAGAGCCCGACGAACACGCGCTTGTCGTCTACTGACGCGCGCCCCCACCACGACTT
AGGGCCGGTCGGGTCAGCGTCGGCGTCTTCGGGGAACCATTGGTCAAGGGGAAGCTTCGGGG
CTTCGGCGGCTTCAAGTTCGGCAAGCCGCTCTTCCGCCCCTTGCTGCCGGAGCGTCAGCGCT
GCCTGTTGCTTCCGGAAGTGCTTCCTGCCAACGGGTCCGTCGTACGCGCCTGCCGCGCGGTC
TTCGTACAGCTCTTCAAGGGCGTTCAGGGCGTCGGCGCGCTCCGCAACAAGGTTCGCCCGTT
CGCCGCTCTTCTCAGGCGCCTCAGTGAGCTTGCCGAAGCGTCGGGCGGCTTCCCACAGAAGC
GCCAACGTCTCTTCGTCGCCTTCGGCGTGCCTGATCTTGTTGAAGATGCGTTCCGCAACGAA
CTTGTCGAGTGCCGCCATGCTGACGTTGCACGTGCCTTCGTGCTGCCCAGGTGCGGACGGGT
CGACCACCTTCCGGCGACGGCAGCGGTAAGAGTCCTTGATCGATTCTTCCCCGCGCTTCGAA
GTCATGACGGCGCCACACTCGCAGTACAGCTTGTCCATGGCGGACAGAATGGCTTGCCCCCG
GGAAAGCCCCTTGCCGCGCCCCCTGCCGTCCAACCACGCCTGAAGCTCATACCACTCAGCGG
GCTCGATGATCGGTCCGCAATCAAGCTCGACCGGCCGGAGCGTGATCGGGTCGCGCTGAATG
CGGTAACCCTCAATCTTCGTGGTCGGCGTGCCGTCCGGCTTCTTCTTGTAGATCACCTCAGC
GGCGAAGCCCGCAATACGCGGGTCCCGAAGGATTCGCATAACGGTTGCCGGGTCCCAGGCGC
TTGAAGCGGTCTTCTTCCCAATCGTCTCGCCCCGGGTCGGCACGGCGTCAGCGTCCATGCGC
TTACAAAGCCCCGTGATGCTGCCCGGGTGAATGGCGGCTTGACTGCCCGGCTTGAAGGGAAG
GTGTTTGTGCGTCTTGATCTCACGCCACCACCACCGGATTACGTCGGGCTCGAACTCGAAGG
GTCCGGTAAGGGGAGTGGTCGAGTGCGCAAGCTTGTTGATGACGACATTGACCATTCGGCCG
TTGCGCGTGATCTCCTTCGTCTCCGAAACAAGCTCGAAGCCGTAAGGCGCCTTCCCGCCGAC
GTACCCGCCCAATTCGCGCTGAAGGTTCTTCGTGTCGAGAATCTTCGCCGACTTCAGCGAAG
ATTCTTTGTGCGACGCGTCGAGCCGCATAATCAGGTGAATCAGGTCCATGACGTTTCCCTGC
CGGAAGACGCCTTCCTGAGTGGAAACAATCGTCACGCCCAGGGCGAGCAATTCCGAGACAAT
CGGAATCGCGTCCATGACCTTCAGGCGCGAGAAGCGCGACACGTCATAGACAATGATCATGT
TGAGCCGCCCGGCGCGGCATTCGTTCAGGATGCGTTCGAACTCCGGGCGCTCCGCCGTCCCG
AACGCCGACGTGCCCGGCGCTTCGCTGAAATGCCCGACGAACCTGAACCGGCCCCCGTCGCG
CTCGACTTCGCGCTGAAGGTCGGCCGCCTTGTCTTCGTTGGCGCTACGCTGTGTCGCTGGGC
TTGCTGCGCTCGAATTCTCGCGCTCGCGCGACTGACGGTCGTAAGCACCCGCGTACGTGTCC
ACCCCGGTCACAACCCCTTGTGTCATGTCGGCGACCCTACGACTAGTGAGCTCGTCGACCCG
GGAATTCCGGACCGGTACCTGCAGGCGTACCTTCTATAGTGTCACCTAAATAGCTTTTTGCA
AAAGCCTAGGCTAGAGTCCGGAGGCTGGATCGGTCCCGGTGTCTTCTATGGAGGTCAAAACA
GCGTGGATGGCGTCTCCAGGCGATCTGACGGTTCACTAAACGAGCTCTGCTTATATAGACCT
CCCACCGTACACGCCTACCGCCCATTTGCGTCAATGGGCGGAGTTGTTACGACATTTTGGA
AAGTCCCGTTGATTTTGGTGCCAAAACAAACTCCCATTGACGTCAATGGGGTGGAGACTTGG
AAATCCCCGTGAGTCAAACCGCTATCCACGCCCATTGATGTACTGCCAAAACCGCATCACCA
TGGTAATAGCGATGACTAATACGTAGATGTACTGCCAAGTAGGAAAGTCCCATAAGGTCATG
TACTGGGCATAATGCCAGGCGGGCCATTTACCGTCATTGACGTCAATAGGGGGCGTACTTGG
CATATGATACACTTGATGTACTGCCAAGTGGGCAGTTTACCGTAAATACTCCACCCATTGAC
```

```
GTCAATGGAAAGTCCCTATTGGCGTTACTATGGGAACATACGTCATTATTGACGTCAATGGG
CGGGGGTCGTTGGGCGGTCAGCCAGGCGGGCCATTTACCGTAAGTTATGTAACGACCTGCAC
GATGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACATTATACGAGCCGGAA
GCTATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAAAGGGCCTCGTATACGCCTATTTTT
ATAGGTTAATGTCATGATAATAATGGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATG
TGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGA
CAATAACCCTGATAAATGCTTCAATAATATTGAAAAACGCGCGAATTGCAAGCTCTGCATTA
ATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGC
TCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCG
GTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCA
GCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCC
CTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAA
AGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCT
TACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCT
GTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCC
GTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA
CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCG
GTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGT
ATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAA
ACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAA
AAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAAC
TCACGTTAAGGGATTTTGGTCATGCCATAACTTCGTATAGCATACATTATACGAAGTTATGG
CATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAAT
CAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA
CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGAT
AACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCAC
GCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGT
GGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAG
TAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCAC
GCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGA
TCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAA
GTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGC
CATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGT
ATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAG
AACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTAC
CGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTT
ACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAAT
AAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTT
ATCAGGGTTATTGTCTCATGCCAGGGGTGGGCACACATATTTGATACCAGCGATCCCTACAC
AGCACATAATTCAATGCGACTTCCCTCTATCGCACATCTTAGACCTTTATTCTCCTCCAGC
ACACATCGAAGCTGCCGAGCAAGCCGTTCTCACCAGTCCAAGACCTGGCATGAGCGGATACA
TATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG
CCACCTGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAG
CTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCG
AGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCC
AACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTA
ATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCC
GATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAA
GGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGC
CGCGCTTAATGCGCCGCTACAGGGCGCGTC
```

Fig. 9

pCMV-luc-attB (SEQ ID NO: 2)

```
CTCTATCGATAGGTACCGAGCTCTTACGCGTGCTAGCCCTCGAGCAGGATCTATACATTGAA
TCAATATTGGCAATTAGCCATATTAGTCATTGGTTATATAGCATAAATCAATATTGGCTATT
GGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATA
TGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATT
AGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCT
GACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCA
ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTA
TTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCG
GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC
ACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGC
GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGC
CTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCC
CCTCGAAGCTCGACTCTAGGGGCTCGAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTAC
TGTTGGTAAAGCCACCATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTATC
CGCTGGAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCCTGGTT
CCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGGACATCACTTACGCTGAGTACTT
CGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCTGAATACAAATCACAGAA
TCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCCGGTGTTGGGCGCGTTATTTATC
GGAGTTGCAGTTGCGCCCGCGAACGACATTTATAATGAACGTGAATTGCTCAACAGTATGGG
CATTTCGCAGCCTACCGTGGTGTTCGTTTCCAAAAAGGGGTTGCAAAAAATTTTGAACGTGC
AAAAAAAGCTCCCAATCATCCAAAAAATTATTATCATGGATTCTAAAACGGATTACCAGGGA
TTTCAGTCGATGTACACGTTCGTCACATCTCATCTACCTCCCGGTTTTAATGAATACGATTT
TGTGCCAGAGTCCTTCGATAGGGACAAGACAATTGCACTGATCATGAACTCCTCTGGATCTA
CTGGTCTGCCTAAAGGTGTCGCTCTGCCTCATAGAACTGCCTGCGTGAGATTCTCGCATGCC
AGAGATCCTATTTTTGGCAATCAAATCATTCCGGATACTGCGATTTTAAGTGTTGTTCCATT
CCATCACGGTTTTGGAATGTTTACTACACTCGGATATTTGATATGTGGATTTCGAGTCGTCT
TAATGTATAGATTTGAAGAAGAGCTGTTTCTGAGGAGCCTTCAGGATTACAAGATTCAAAGT
GCGCTGCTGGTGCCAACCCTATTCTCCTTCTTCGCCAAAAGCACTCTGATTGACAAATACGA
TTTATCTAATTTACACGAAATTGCTTCTGGTGGCGCTCCCCTCTCTAAGGAAGTCGGGGAAG
CGGTTGCCAAGAGGTTCCATCTGCCAGGTATCAGGCAAGGATATGGGCTCACTGAGACTACA
TCAGCTATTCTGATTACACCCGAGGGGGATGATAAACCGGGCGCGGTCGGTAAAGTTGTTCC
ATTTTTTGAAGCGAAGGTTGTGGATCTGGATACCGGGAAAACGCTGGGCGTTAATCAAAGAG
GCGAACTGTGTGTGAGAGGTCCTATGATTATGTCCGGTTATGTAAACAATCCGGAAGCGACC
AACGCCTTGATTGACAAGGATGGATGGCTACATTCTGGAGACATAGCTTACTGGGACGAAGA
CGAACACTTCTTCATCGTTGACCGCCTGAAGTCTCTGATTAAGTACAAAGGCTATCAGGTGG
CTCCCGCTGAATTGGAATCCATCTTGCTCCAACACCCCAACATCTTCGACGCAGGTGTCGCA
GGTCTTCCCGACGATGACGCCGGTGAACTTCCCGCCGCCGTTGTTGTTTTGGAGCACGGAAA
GACGATGACGGAAAAAGAGATCGTGGATTACGTCGCCAGTCAAGTAACAACCGCGAAAAAGT
TGCGCGGAGGAGTTGTGTTTGTGGACGAAGTACCGAAAGGTCTTACCGGAAAACTCGACGCA
AGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGGAAAGATCGCCGTGTAATTCTA
GAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAA
CCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTA
TTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTT
TCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTA
AAATCGATAAGGATCAATTCGGCTTCAGGTACCGTCGACGATGTAGGTCACGGTCTCGAAGC
CGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACCTCACCCATC
TGGTCCATCATGATGAACGGGTCGAGGTGGCGGTAGTTGATCCCGGCGAACGCGCGGCGCAC
```

```
CGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGTGAGCACGGGACGTGCGA
CGGCGTCGGCGGGTGCGGATACGCGGGGCAGCGTCAGCGGGTTCTCGACGGTCACGGCGGGC
ATGTCGACAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTC
CTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGC
AACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC
GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAG
AATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGT
AAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAA
TCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCC
CTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCC
TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGT
GTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCG
CCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCA
GCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAA
GTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGC
CAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGC
GGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCC
TTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGG
TCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAA
TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC
ACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGA
TAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCA
CGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAG
TGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAA
GTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCA
CGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATG
ATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTA
AGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATG
CCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTG
TATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCA
GAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTA
CCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTT
TACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAA
TAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATT
TATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAAT
AGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGCATTAA
GCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCC
GCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCT
AAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAAC
TTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTG
ACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCC
TATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAA
ATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTCC
CATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATT
ACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTTGGAGCGG
CCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAG
TACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTC
CCCAGTGCAAGTGCAGGTGCCAGAACATTT
```

Fig. 10

pCMV-luc-attP (SEQ ID NO: 3)

CTCTATCGATAGGTACCGAGCTCTTACGCGTGCTAGCCCTCGAGCAGGATCTATACATTGAA
TCAATATTGGCAATTAGCCATATTAGTCATTGGTTATATAGCATAAATCAATATTGGCTATT
GGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATA
TGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATT
AGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCT
GACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCA
ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTA
TTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCG
GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC
ACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGC
GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGC
CTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCC
CCTCGAAGCTCGACTCTAGGGGCTCGAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTAC
TGTTGGTAAAGCCACCATGGAAGACGCCAAAAACATAAAGAAAGGCCCGGCGCCATTCTATC
CGCTGGAAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCCTGGTT
CCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGGACATCACTTACGCTGAGTACTT
CGAAATGTCCGTTCGGTTGGCAGAAGCTATGAAACGATATGGGCTGAATACAAATCACAGAA
TCGTCGTATGCAGTGAAAACTCTCTTCAATTCTTTATGCCGGTGTTGGGCGCGTTATTTATC
GGAGTTGCAGTTGCGCCCGCGAACGACATTTATAATGAACGTGAATTGCTCAACAGTATGGG
CATTTCGCAGCCTACCGTGGTGTTCGTTTCCAAAAAGGGGTTGCAAAAAATTTTGAACGTGC
AAAAAAAGCTCCCAATCATCCAAAAAATTATTATCATGGATTCTAAAACGGATTACCAGGGA
TTTCAGTCGATGTACACGTTCGTCACATCTCATCTACCTCCCGGTTTTAATGAATACGATTT
TGTGCCAGAGTCCTTCGATAGGGACAAGACAATTGCACTGATCATGAACTCCTCTGGATCTA
CTGGTCTGCCTAAAGGTGTCGCTCTGCCTCATAGAACTGCCTGCGTGAGATTCTCGCATGCC
AGAGATCCTATTTTTGGCAATCAAATCATTCCGGATACTGCGATTTTAAGTGTTGTTCCATT
CCATCACGGTTTTGGAATGTTTACTACACTCGGATATTTGATATGTGGATTTCGAGTCGTCT
TAATGTATAGATTTGAAGAAGAGCTGTTTCTGAGGAGCCTTCAGGATTACAAGATTCAAAGT
GCGCTGCTGGTGCCAACCCTATTCTCCTTCTTCGCCAAAAGCACTCTGATTGACAAATACGA
TTTATCTAATTTACACGAAATTGCTTCTGGTGGCGCTCCCCTCTCTAAGGAAGTCGGGGAAG
CGGTTGCCAAGAGGTTCCATCTGCCAGGTATCAGGCAAGGATATGGGCTCACTGAGACTACA
TCAGCTATTCTGATTACACCCGAGGGGGATGATAAACCGGGCGCGGTCGGTAAAGTTGTTCC
ATTTTTTGAAGCGAAGGTTGTGGATCTGGATACCGGGAAAACGCTGGGCGTTAATCAAAGAG
GCGAACTGTGTGTGAGAGGTCCTATGATTATGTCCGGTTATGTAAACAATCCGGAAGCGACC
AACGCCTTGATTGACAAGGATGGATGGCTACATTCTGGAGACATAGCTTACTGGGACGAAGA
CGAACACTTCTTCATCGTTGACCGCCTGAAGTCTCTGATTAAGTACAAAGGCTATCAGGTGG
CTCCCGCTGAATTGGAATCCATCTTGCTCCAACACCCCAACATCTTCGACGCAGGTGTCGCA
GGTCTTCCCGACGATGACGCCGGTGAACTTCCCGCCGCCGTTGTTGTTTTGGAGCACGGAAA
GACGATGACGGAAAAAGAGATCGTGGATTACGTCGCCAGTCAAGTAACAACCGCGAAAAAGT
TGCGCGGAGGAGTTGTGTTTGTGGACGAAGTACCGAAAGGTCTTACCGGAAAACTCGACGCA
AGAAAAATCAGAGAGATCCTCATAAAGGCCAAGAAGGGCGGAAAGATCGCCGTGTAATTCTA
GAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAA
CCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTA
TTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTT
TCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTA
AAATCGATAAGGATCAATTCGGCTTCGACTAGTACTGACGGACACACCGAAGCCCCGGCGGC
AACCCTCAGCGGATGCCCCGGGGCTTCACGTTTTCCCAGGTCAGAAGCGGTTTTCGGGAGTA
GTGCCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGGCGTAGGGTCGCCGACATGAC
ACAAGGGGTTGTGACCGGGGGTGGACACGTACGCGGGTGCTTACGACCGTCAGTCGCGCGAGC

```
GCGACTAGTACAAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCA
GCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATC
ATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGC
GCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCC
ACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA
CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACA
AAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTT
CCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTC
CGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTT
CGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGC
TGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACT
GGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCT
TGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTG
AAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGG
TAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAG
ATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATT
TTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTT
TAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTG
AGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTG
TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCA
GAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGA
GTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGT
GTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTA
CATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGA
AGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGT
CATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAAT
AGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACAT
AGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGAT
CTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCAT
CTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAG
GGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAG
CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAAC
AAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGCA
TTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGC
GCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAG
CTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAA
AAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCC
TTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCA
ACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTA
AAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAAT
TTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGC
TATTACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTTGGA
GCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCG
ATAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGC
TGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTT
```

Fig. 11

pCMV-pur-attB (SEQ ID NO: 4)

```
CTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGAC
AAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCT
TTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTAT
GTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTG
GTAAAATCGATAAGGATCAATTCGGCTTCAGGTACCGTCGACGATGTAGGTCACGGTCTCGA
AGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACCTCACCC
ATCTGGTCCATCATGATGAACGGGTCGAGGTGGCGGTAGTTGATCCCGGCGAACGCGCGGCG
CACCGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGTGAGCACGGGACGTG
CGACGGCGTCGGCGGGTGCGGATACGCGGGGCAGCGTCAGCGGGTTCTCGACGGTCACGGCG
GGCATGTCGACAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAG
CTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCA
TGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCG
CTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA
CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC
CGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAA
AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC
CCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC
GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTC
GGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT
GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT
GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGA
AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGT
AGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAGGATCTCAAGAAGA
TCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTT
TGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTT
AAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA
GGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGT
AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAG
TAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTG
TCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTAC
ATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAA
GTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATA
GTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATA
GCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATC
TTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATC
TTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGG
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGC
ATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACA
AATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGCAT
TAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCG
CCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGC
TCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAA
AACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCT
TTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAA
CCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAA
AAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATT
```

```
TCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT
ATTACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTTGGAG
CGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTGTTGGTTTTTTTGTGTGAATCGA
TAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCT
GTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGATAGGTACCGAGCTCTTACGC
GTGCTAGCCCTCGAGCAGGATCTATACATTGAATCAATATTGGCAATTAGCCATATTAGTCA
TTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCAT
AATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGAC
TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCG
TTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACG
TCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGT
GGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC
CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA
CGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG
GTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC
ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT
CGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATAT
AAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACC
TCCATAGAAGACACCGGGACCGATCCAGCCTCCCCTCGAAGCTCGACTCTAGGGGCTCGAGA
TCTGCGATCTAAGTAAGCTTGCATGCCTGCAGGTCGGCCGCCACGACCGGTGCCGCCACCAT
CCCCTGACCCACGCCCCTGACCCCTCACAAGGAGACGACCTTCCATGACCGAGTACAAGCCC
ACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCACCCTCGCCGCCGCGTT
CGCCGACTACCCCGCCACGCGCCACACCGTCGACCCGGACCGCCACATCGAGCGGGTCACCG
AGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGGAC
GACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCGC
CGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGG
AAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTC
TCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGC
CGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACG
AGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTGGTGC
ATGACCCGCAAGCCCGGTGCCTGACGCCCGCCCCACGACCCGCAGCGCCCGACCGAAAGGAG
CGCACGACCCCATGGCTCCGACCGAAGCCGACCCGGGCGGCCCCGCCGACCCCGCACCCGCC
CCCGAGGCCCACCGACT
```

Fig. 12

pCMV-pur-attP (SEQ ID NO: 5)

```
CTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGAC
AAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCT
TTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTAT
GTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTG
GTAAAATCGATAAGGATCAATTCGGCTTCGACTAGTACTGACGGACACACCGAAGCCCCGGC
GGCAACCCTCAGCGGATGCCCCGGGGCTTCACGTTTTCCCAGGTCAGAAGCGGTTTTCGGGA
GTAGTGCCCCAACTGGGGTAACCTTTGAGTTCTCTCAGTTGGGGGCGTAGGGTCGCCGACAT
GACACAAGGGGTTGTGACCGGGGTGGACACGTACGCGGGTGCTTACGACCGTCAGTCGCGCG
AGCGCGACTAGTACAAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAG
TCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTT
ATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCCTCGCTCACTGACTCGC
TGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTA
TCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAG
GAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATC
ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCT
GTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCA
GTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGAC
CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCC
ACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGT
TCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTG
CTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGC
TGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAG
AAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGG
ATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAG
TTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCA
GTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTC
GTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCG
AGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGC
GCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT
AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGT
GGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAG
TTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTC
AGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTAC
TGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAG
AATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCA
CATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAG
GATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAG
CATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAA
AAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTG
AAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATA
AACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGC
GCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCT
AGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTC
AAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCC
AAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCG
CCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACAC
TCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGG
TTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTAC
AATTTCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTT
```

CGCTATTACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTT
GGAGCGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTGAA
TCGATAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATA
GGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGATAGGTACCGAGCTCTT
ACGCGTGCTAGCCCTCGAGCAGGATCTATACATTGAATCAATATTGGCAATTAGCCATATTA
GTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATA
TCATAATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTAT
TGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTC
CGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATT
GACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT
GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGT
CCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGAC
CTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGA
TGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGT
CTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAA
ATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCT
ATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTT
GACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCCCTCGAAGCTCGACTCTAGGGGCTC
GAGATCTGCGATCTAAGTAAGCTTGCATGCCTGCAGGTCGGCCGCCACGACCGGTGCCGCCA
CCATCCCCTGACCCACGCCCCTGACCCCTCACAAGGAGACGACCTTCCATGACCGAGTACAA
GCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCACCCTCGCCGCCG
CGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGACCCGGACCGCCACATCGAGCGGGTC
ACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGC
GGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGT
TCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAG
ATGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGG
CGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGG
CGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTC
TACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTG
GTGCATGACCCGCAAGCCCGGTGCCTGACGCCCGCCCCACGACCCGCAGCGCCCGACCGAAA
GGAGCGCACGACCCCATGGCTCCGACCGAAGCCGACCCGGGCGGCCCCGCCGACCCCGCACC
CGCCCCCGAGGCCCACCGACT

Fig. 13

pCMV-EGFP-attB (SEQ ID NO: 6)

```
CTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGAC
AAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCT
TTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTAT
GTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTG
GTAAAATCGATAAGGATCAATTCGGCTTCAGGTACCGTCGACGATGTAGGTCACGGTCTCGA
AGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACCTCACCC
ATCTGGTCCATCATGATGAACGGGTCGAGGTGGCGGTAGTTGATCCCGGCGAACGCGCGGCG
CACCGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGTGAGCACGGGACGTG
CGACGGCGTCGGCGGGTGCGGATACGCGGGGCAGCGTCAGCGGGTTCTCGACGGTCACGGCG
GGCATGTCGACAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAG
CTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCA
TGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCG
CTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCA
CAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAAC
CGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAA
AAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTC
CCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC
GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTC
GGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCT
GCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTT
GAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGA
AGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGT
AGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGA
TCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTT
TGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTT
AAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGA
GGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGT
AGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAG
AAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAG
TAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTG
TCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTAC
ATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAA
GTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTC
ATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATA
GTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATA
GCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATC
TTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATC
TTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGG
GAATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGC
ATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACA
AATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGCAT
TAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCG
CCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGC
TCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAA
AACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCT
TTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAA
CCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAA
AAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATT
```

TCCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCT
ATTACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTTGGAG
CGGCCGCAATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTTGTGTGAATCGA
TAGTACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCT
GTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGATAGGTACCGAGCTCTTACGC
GTGCTAGCCCTCGAGCAGGATCTATACATTGAATCAATATTGGCAATTAGCCATATTAGTCA
TTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCAT
AATATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGAC
TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCG
TTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACG
TCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGT
GGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC
CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA
CGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG
GTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC
ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT
CGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATAT
AAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACC
TCCATAGAAGACACCGGGACCGATCCAGCCTCCCCTCGAAGCTCGACTCTAGGGGCTCGAGA
TCCCCGGGTACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGC
CCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGC
GAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCC
CGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACC
CCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAG
CGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGG
CGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCC
TGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAG
AAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCT
CGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACC
ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC
CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAG
CGGCCGCTCGAGCATGCAT

Fig. 14

p-12.0-lys-LSPIFNMM-CMV-pur-attB (SEQ ID NO: 7)

```
GGGCTGCAGGAATTCGATTGCCGCCTTCTTTGATATTCACTCTGTTGTATTTCATCTCTTCT
TGCCGATGAAAGGATATAACAGTCTGTATAACAGTCTGTGAGGAAATACTTGGTATTTCTTC
TGATCAGTGTTTTTATAAGTAATGTTGAATATTGGATAAGGCTGTGTGTCCTTTGTCTTGGG
AGACAAAGCCCACAGCAGGTGGTGGTTGGGGTGGTGGCAGCTCAGTGACAGGAGAGGTTTTT
TTGCCTGTTTTTTTTTTTTTTTTTTTTTTAAGTAAGGTGTTCTTTTTTCTTAGTAAATTTT
CTACTGGACTGTATGTTTTGACAGGTCAGAAACATTTCTTCAAAAGAAGAACCTTTTGGAAA
CTGTACAGCCCTTTTCTTTCATTCCCTTTTTGCTTTCTGTGCCAATGCCTTTGGTTCTGATT
GCATTATGGAAAACGTTGATCGGAACTTGAGGTTTTTATTTATAGTGTGGCTTGAAAGCTTG
GATAGCTGTTGTTACACGAGATACCTTATTAAGTTTAGGCCAGCTTGATGCTTTATTTTTTC
CCTTTGAAGTAGTGAGCGTTCTCTGGTTTTTTTCCTTTGAAACTGGTGAGGCTTAGATTTTT
CTAATGGGATTTTTTACCTGATGATCTAGTTGCATACCCAAATGCTTGTAAATGTTTTCCTA
GTTAACATGTTGATAACTTCGGATTTACATGTTGTATATACTTGTCATCTGTGTTTCTAGTA
AAAATATATGGCATTTATAGAAATACGTAATTCCTGATTTCCTTTTTTTTTATCTCTATGCT
CTGTGTGTACAGGTCAAACAGACTTCACTCCTATTTTTATTTATAGAATTTTATATGCAGTC
TGTCGTTGGTTCTTGTGTTGTAAGGATACAGCCTTAAATTTCCTAGAGCGATGCTCAGTAAG
GCGGGTTGTCACATGGGTTCAAATGTAAAACGGGCACGTTTGGCTGCTGCCTTCCCGAGATC
CAGGACACTAAACTGCTTCTGCACTGAGGTATAAATCGCTTCAGATCCCAGGGAAGTGCAGA
TCCACGTGCATATTCTTAAAGAAGAATGAATACTTTCTAAAATATTTTGGCATAGGAAGCAA
GCTGCATGGATTTGTTTGGGACTTAAATTATTTTGGTAACGGAGTGCATAGGTTTTAAACAC
AGTTGCAGCATGCTAACGAGTCACAGCGTTTATGCAGAAGTGATGCCTGGATGCCTGTTGCA
GCTGTTTACGGCACTGCCTTGCAGTGAGCATTGCAGATAGGGGTGGGGTGCTTTGTGTCGTG
TTCCCACACGCTGCCACACAGCCACCTCCCGGAACACATCTCACCTGCTGGGTACTTTTCAA
ACCATCTTAGCAGTAGTAGATGAGTTACTATGAAACAGAGAAGTTCCTCAGTTGGATATTCT
CATGGGATGTCTTTTTTCCCATGTTGGGCAAAGTATGATAAAGCATCTCTATTTGTAAATTA
TGCACTTGTTAGTTCCTGAATCCTTTCTATAGCACCACTTATTGCAGCAGGTGTAGGCTCTG
GTGTGGCCTGTGTCTGTGCTTCAATCTTTTAAAGCTTCTTTGGAAATACACTGACTTGATTG
AAGTCTCTTGAAGATAGTAAACAGTACTTACCTTTGATCCCAATGAAATCGAGCATTTCAGT
TGTAAAAGAATTCCGCCTATTCATACCATGTAATGTAATTTTACACCCCCAGTGCTGACACT
TTGGAATATATTCAAGTAATAGACTTTGGCCTCACCCTCTTGTGTACTGTATTTTGTAATAG
AAAATATTTTAAACTGTGCATATGATTATTACATTATGAAAGAGACATTCTGCTGATCTTCA
AATGTAAGAAAATGAGGAGTGCGTGTGCTTTTATAAATACAAGTGATTGCAAATTAGTGCAG
GTGTCCTTAAAAAAAAAAAAAAAAAAAGTAATATAAAAAGGACCAGGTGTTTTACAAGTGAAAT
ACATTCCTATTTGGTAAACAGTTACATTTTTATGAAGATTACCAGCGCTGCTGACTTTCTAA
ACATAAGGCTGTATTGTCTTCCTGTACCATTGCATTTCCTCATTCCCAATTTGCACAAGGAT
GTCTGGGTAAACTATTCAAGAAATGGCTTTGAAATACAGCATGGGAGCTTGTCTGAGTTGGA
ATGCAGAGTTGCACTGCAAAATGTCAGGAAATGGATGTCTCTCAGAATGCCCAACTCCAAAG
GATTTTATATGTGTATATAGTAAGCAGTTTCCTGATTCCAGCAGGCCAAAGAGTCTGCTGAA
TGTTGTGTTGCCGGAGACCTGTATTTCTCAACAAGGTAAGATGGTATCCTAGCAACTGCGGA
TTTTAATACATTTTCAGCAGAAGTACTTAGTTAATCTCTACCTTTAGGGATCGTTTCATCAT
TTTTAGATGTTATACTTGAAATACTGCATAACTTTTAGCTTTCATGGGTTCCTTTTTTTCAG
CCTTTAGGAGACTGTTAAGCAATTTGCTGTCCAACTTTTGTGTTGGTCTTAAACTGCAATAG
TAGTTTACCTTGTATTGAAGAAATAAAGACCATTTTTATATTAAAAAATACTTTTGTCTGTC
TTCATTTTGACTTGTCTGATATCCTTGCAGTGCCCATTATGTCAGTTCTGTCAGATATTCAG
ACATCAAAACTTAACGTGAGCTCAGTGGAGTTACAGCTGCGGTTTTGATGCTGTTATTATTT
CTGAAACTAGAAATGATGTTGTCTTCATCTGCTCATCAAACACTTCATGCAGAGTGTAAGGC
TAGTGAGAAATGCATACATTTATTGATACTTTTTTAAAGTCAACTTTTTATCAGATTTTTTT
TTCATTTGGAAATATATTGTTTTCTAGACTGCATAGCTTCTGAATCTGAAATGCAGTCTGAT
TGGCATGAAGAAGCACAGCACTCTTCATCTTACTTAAACTTCATTTTGGAATGAAGGAAGTT
AAGCAAGGGCACAGGTCCATGAAATAGAGACAGTGCGCTCAGGAGAAAGTGAACCTGGATTT
CTTTGGCTAGTGTTCTAAATCTGTAGTGAGGAAAGTAACACCCGATTCCTTGAAAGGGCTCC
```

```
AGCTTTAATGCTTCCAAATTGAAGGTGGCAGGCAACTTGGCCACTGGTTATTTACTGCATTA
TGTCTCAGTTTCGCAGCTAACCTGGCTTCTCCACTATTGAGCATGGACTATAGCCTGGCTTC
AGAGGCCAGGTGAAGGTTGGGATGGGTGGAAGGAGTGCTGGGCTGTGGCTGGGGGGACTGTG
GGGACTCCAAGCTGAGCTTGGGGTGGGCAGCACAGGGAAAAGTGTGGGTAACTATTTTTAAG
TACTGTGTTGCAAACGTCTCATCTGCAAATACGTAGGGTGTGTACTCTCGAAGATTAACAGT
GTGGGTTCAGTAATATATGGATGAATTCACAGTGGAAGCATTCAAGGGTAGATCATCTAACG
ACACCAGATCATCAAGCTATGATTGGAAGCGGTATCAGAAGAGCGAGGAAGGTAAGCAGTCT
TCATATGTTTTCCCTCCACGTAAAGCAGTCTGGGAAAGTAGCACCCCTTGAGCAGAGACAAG
GAAATAATTCAGGAGCATGTGCTAGGAGAACTTTCTTGCTGAATTCTACTTGCAAGAGCTTT
GATGCCTGGCTTCTGGTGCCTTCTGCAGCACCTGCAAGGCCCAGAGCCTGTGGTGAGCTGGA
GGGAAAGATTCTGCTCAAGTCCAAGCTTCAGCAGGTCATTGTCTTTGCTTCTTCCCCCAGCA
CTGTGCAGCAGAGTGGAACTGATGTCGAAGCCTCCTGTCCACTACCTGTTGCTGCAGGCAGA
CTGCTCTCAGAAAAAGAGAGCTAACTCTATGCCATAGTCTGAAGGTAAAATGGGTTTTAAAA
AAGAAAACACAAAGGCAAAACCGGCTGCCCCATGAGAAGAAAGCAGTGGTAAACATGGTAGA
AAAGGTGCAGAAGCCCCCAGGCAGTGTGACAGGCCCCTCCTGCCACCTAGAGGCGGGAACAA
GCTTCCCTGCCTAGGGCTCTGCCCGCGAAGTGCGTGTTTCTTTGGTGGGTTTTGTTTGGCGT
TTGGTTTTGAGATTTAGACACAAGGGAAGCCTGAAAGGAGGTGTTGGGCACTATTTTGGTTT
GTAAAGCCTGTACTTCAAATATATATTTTGTGAGGGAGTGTAGCGAATTGGCCAATTTAAAA
TAAAGTTGCAAGAGATTGAAGGCTGAGTAGTTGAGAGGGTAACACGTTTAATGAGATCTTCT
GAAACTACTGCTTCTAAACACTTGTTTGAGTGGTGAGACCTTGGATAGGTGAGTGCTCTTGT
TACATGTCTGATGCACTTGCTTGTCCTTTTCCATCCACATCCATGCATTCCACATCCACGCA
TTTGTCACTTATCCCATATCTGTCATATCTGACATACCTGTCTCTTCGTCACTTGGTCAGAA
GAAACAGATGTGATAATCCCCAGCCGCCCCAAGTTTGAGAAGATGGCAGTTGCTTCTTTCCC
TTTTTCCTGCTAAGTAAGGATTTTCTCCTGGCTTTGACACCTCACGAAATAGTCTTCCTGCC
TTACATTCTGGGCATTATTTCAAATATCTTTGGAGTGCGCTGCTCTCAAGTTTGTGTCTTCC
TACTCTTAGAGTGAATGCTCTTAGAGTGAAAGAGAAGGAAGAGAAGATGTTGGCCGCAGTTC
TCTGATGAACACACCTCTGAATAATGGCCAAAGGTGGGTGGGTTTCTCTGAGGAACGGGCAG
CGTTTGCCTCTGAAAGCAAGGAGCTCTGCGGAGTTGCAGTTATTTTGCAACTGATGGTGGAA
CTGGTGCTTAAAGCAGATTCCCTAGGTTCCCTGCTACTTCTTTTTCCTTCTTGGCAGTCAGTT
TATTTCTGACAGACAAACAGCCACCCCCACTGCAGGCTTAGAAAGTATGTGGCTCTGCCTGG
GTGTGTTACAGCTCTGCCCTGGTGAAAGGGGATTAAAACGGGCACCATTCATCCCAAACAGG
ATCCTCATTCATGGATCAAGCTGTAAGGAACTTGGGCTCCAACCTCAAAACATTAATTGGAG
TACGAATGTAATTAAAACTGCATTCTCGCATTCCTAAGTCATTTAGTCTGGACTCTGCAGCA
TGTAGGTCGGCAGCTCCCACTTTCTCAAAGACCACTGATGGAGGAGTAGTAAAAATGGAGAC
CGATTCAGAACAACCAACGGAGTGTTGCCGAAGAAACTGATGGAAATAATGCATGAATTGTG
TGGTGGACATTTTTTTTAAATACATAAACTACTTCAAATGAGGTCGGAGAAGGTCAGTGTTT
TATTAGCAGCCATAAAACCAGGTGAGCGAGTACCATTTTTCTCTACAAGAAAAACGATTCTG
AGCTCTGCGTAAGTATAAGTTCTCCATAGCGGCTGAAGCTCCCCCCTGGCTGCCTGCCATCT
CAGCTGGAGTGCAGTGCCATTTCCTTGGGGTTTCTCTCACAGCAGTAATGGGACAATACTTC
ACAAAAATTCTTTCTTTTCCTGTCATGTGGGATCCCTACTGTGCCCTCCTGGTTTTACGTTA
CCCCCTGACTGTTCCATTCAGCGGTTTGGAAAGAGAAAAGAATTTGGAAATAAAACATGTC
TACGTTATCACCTCCTCCAGCATTTTGGTTTTTAATTATGTCAATAACTGGCTTAGATTTGG
AAATGAGAGGGGGTTGGGTGTATTACCGAGGAACAAAGGAAGGCTTATATAAACTCAAGTCT
TTTATTTAGAGAACTGGCAAGCTGTCAAAAACAAAAAGGCCTTACCACCAAATTAAGTGAAT
AGCCGCTATAGCCAGCAGGGCCAGCACGAGGGATGGTGCACTGCTGGCACTATGCCACGGCC
TGCTTGTGACTCTGAGAGCAACTGCTTTGGAAATGACAGCACTTGGTGCAATTTCCTTTGTT
TCAGAATGCGTAGAGCGTGTGCTTGGCGACAGTTTTTCTAGTTAGGCCACTTCTTTTTTCCT
TCTCTCCTCATTCTCCTAAGCATGTCTCCATGCTGGTAATCCCAGTCAAGTGAACGTTCAAA
CAATGAATCCATCACTGTAGGATTCTCGTGGTGATCAAATCTTTGTGTGAGGTCTATAAAAT
ATGGAAGCTTATTTATTTTTCGTTCTTCCATATCAGTCTTCTCTATGACAATTCACATCCAC
CACAGCAAATTAAAGGTGAAGGAGGCTGGTGGGATGAAGAGGGTCTTCTAGCTTTACGTTCT
TCCTTGCAAGGCCACAGGAAAATGCTGAGAGCTGTAGAATACAGCCTGGGGTAAGAAGTTCA
```

GTCTCCTGCTGGGACAGCTAACCGCATCTTATAACCCCTTCTGAGACTCATCTTAGGACCAA
ATAGGGTCTATCTGGGGTTTTTGTTCCTGCTGTTCCTCCTGGAAGGCTATCTCACTATTTCA
CTGCTCCCACGGTTACAAACCAAAGATACAGCCTGAATTTTTTCTAGGCCACATTACATAAA
TTTGACCTGGTACCAATATTGTTCTCTATATAGTTATTTCCTTCCCCACTGTGTTTAACCCC
TTAAGGCATTCAGAACAACTAGAATCATAGAATGGTTTGGATTGGAAGGGGCCTTAAACATC
ATCCATTTCCAACCCTCTGCCATGGGCTGCTTGCCACCCACTGGCTCAGGCTGCCCAGGGCC
CCATCCAGCCTGGCCTTGAGCACCTCCAGGGATGGGGCACCCACAGCTTCTCTGGGCAGCCT
GTGCCAACACCTCACCACTCTCTGGGTAAAGAATTCTCTTTTAACATCTAATCTAAATCTCT
TCTCTTTTAGTTTAAAGCCATTCCTCTTTTTCCCGTTGCTATCTGTCCAAGAAATGTGTATT
GGTCTCCCTCCTGCTTATAAGCAGGAAGTACTGGAAGGCTGCAGTGAGGTCTCCCCACAGCC
TTCTCTTCTCCAGGCTGAACAAGCCCAGCTCCTTCAGCCTGTCTTCGTAGGAGATCATCTTA
GTGGCCCTCCTCTGGACCCATTCCAACAGTTCCACGGCTTTCTTGTGGAGCCCCAGGTCTGG
ATGCAGTACTTCAGATGGGGCCTTACAAAGGCAGAGCAGATGGGGACAATCGCTTACCCCTC
CCTGCTGGCTGCCCCTGTTTTGATGCAGCCCAGGGTACTGTTGGCCTTTCAGGCTCCCAGAC
CCCTTGCTGATTTGTGTCAAGCTTTTCATCCACCAGAACCCACGCTTCCTGGTTAATACTTC
TGCCCTCACTTCTGTAAGCTTGTTTCAGGAGACTTCCATTCTTTAGGACAGACTGTGTTACA
CCTACCTGCCCTATTCTTGCATATATACATTTCAGTTCATGTTTCCTGTAACAGGACAGAAT
ATGTATTCCTCTAACAAAAATACATGCAGAATTCCTAGTGCCATCTCAGTAGGGTTTTCATG
GCAGTATTAGCACATAGTCAATTTGCTGCAAGTACCTTCCAAGCTGCGGCCTCCCATAAATC
CTGTATTTGGGATCAGTTACCTTTTGGGGTAAGCTTTTGTATCTGCAGAGACCCTGGGGGTT
CTGATGTGCTTCAGCTCTGCTCTGTTCTGACTGCACCATTTTCTAGATCACCCAGTTGTTCC
TGTACAACTTCCTTGTCCTCCATCCTTTCCCAGCTTGTATCTTTGACAAATACAGGCCTATT
TTTGTGTTTGCTTCAGCAGCCATTTAATTCTTCAGTGTCATCTTGTTCTGTTGATGCCACTG
GAACAGGATTTTCAGCAGTCTTGCAAAGAACATCTAGCTGAAAACTTTCTGCCATTCAATAT
TCTTACCAGTTCTTCTTGTTTGAGGTGAGCCATAAATTACTAGAACTTCGTCACTGACAAGT
TTATGCATTTTATTACTTCTATTATGTACTTACTTTGACATAACACAGACACGCACATATTT
TGCTGGGATTTCCACAGTGTCTCTGTGTCCTTCACATGGTTTTACTGTCATACTTCCGTTAT
AACCTTGGCAATCTGCCCAGCTGCCCATCACAAGAAAAGAGATTCCTTTTTTATTACTTCTC
TTCAGCCAATAAACAAAATGTGAGAAGCCCAAACAAGAACTTGTGGGGCAGGCTGCCATCAA
GGGAGAGACAGCTGAAGGGTTGTGTAGCTCAATAGAATTAAGAAATAATAAAGCTGTGTCAG
ACAGTTTTGCCTGATTTATACAGGCACGCCCCAAGCCAGAGAGGCTGTCTGCCAAGGCCACC
TTGCAGTCCTTGGTTTGTAAGATAAGTCATAGGTAACTTTTCTGGTGAATTGCGTGGAGAAT
CATGATGGCAGTTCTTGCTGTTTACTATGGTAAGATGCTAAAATAGGAGACAGCAAAGTAAC
ACTTGCTGCTGTAGGTGCTCTGCTATCCAGACAGCGATGGCACTCGCACACCAAGATGAGGG
ATGCTCCCAGCTGACGGATGCTGGGGCAGTAACAGTGGGTCCCATGCTGCCTGCTCATTAGC
ATCACCTCAGCCCTCACCAGCCCATCAGAAGGATCATCCCAAGCTGAGGAAAGTTGCTCATC
TTCTTCACATCATCAAACCTTTGGCCTGACTGATGCCTCCCGGATGCTTAAATGTGGTCACT
GACATCTTTATTTTTCTATGATTTCAAGTCAGAACCTCCGGATCAGGAGGGAACACATAGTG
GGAATGTACCCTCAGCTCCAAGGCCAGATCTTCCTTCAATGATCATGCATGCTACTTAGGAA
GGTGTGTGTGTGTGAATGTAGAATTGCCTTTGTTATTTTTTCTTCCTGCTGTCAGGAACATT
TTGAATACCAGAGAAAAGAAAGTGCTCTTCTTGGCATGGGAGGAGTTGTCACACTTGCAA
AATAAAGGATGCAGTCCCAAATGTTCATAATCTCAGGGTCTGAAGGAGGATCAGAAACTGTG
TATACAATTTCAGGCTTCTCTGAATGCAGCTTTTGAAAGCTGTTCCTGGCCGAGGCAGTACT
AGTCAGAACCCTCGGAAACAGGAACAAATGTCTTCAAGGTGCAGCAGGAGGAAACACCTTGC
CCATCATGAAAGTGAATAACCACTGCCGCTGAAGGAATCCAGCTCCTGTTTGAGCAGGTGCT
GCACACTCCCACACTGAAACAACAGTTCATTTTTATAGGACTTCCAGGAAGGATCTTCTTCT
TAAGCTTCTTAATTATGGTACATCTCCAGTTGGCAGATGACTATGACTACTGACAGGAGAAT
GAGGAACTAGCTGGGAATATTTCTGTTTGACCACCATGGAGTCACCCATTTCTTTACTGGTA
TTTGGAAATAATAATTCTGAATTGCAAAGCAGGAGTTAGCGAAGATCTTCATTTCTTCCATG
TTGGTGACAGCACAGTTCTGGCTATGAAAGTCTGCTTACAAGGAAGAGGATAAAAATCATAG
GGATAATAAATCTAAGTTTGAAGACAATGAGGTTTTAGCTGCATTTGACATGAAGAAATTGA

GACCTCTACTGGATAGCTATGGTATTTACGTGTCTTTTTGCTTAGTTACTTATTGACCCCAG
CTGAGGTCAAGTATGAACTCAGGTCTCTCGGGCTACTGGCATGGATTGATTACATACAACTG
TAATTTTAGCAGTGATTTAGGGTTTATGAGTACTTTTGCAGTAAATCATAGGGTTAGTAATG
TTAATCTCAGGGAAAAAAAAAAAAGCCAACCCTGACAGACATCCCAGCTCAGGTGGAAATC
AAGGATCACAGCTCAGTGCGGTCCCAGAGAACACAGGGACTCTTCTCTTAGGACCTTTATGT
ACAGGGCCTCAAGATAACTGATGTTAGTCAGAAGACTTTCCATTCTGGCCACAGTTCAGCTG
AGGCAATCCTGGAATTTTCTCTCCGCTGCACAGTTCCAGTCATCCCAGTTTGTACAGTTCTG
GCACTTTTTGGGTCAGGCCGTGATCCAAGGAGCAGAAGTTCCAGCTATGGTCAGGGAGTGCC
TGACCGTCCCAACTCACTGCACTCAAACAAAGGCGAAACCACAAGAGTGGCTTTTGTTGAAA
TTGCAGTGTGGCCCAGAGGGGCTGCACCAGTACTGGATTGACCACGAGGCAACATTAATCCT
CAGCAAGTGCAATTTGCAGCCATTAAATTGAACTAACTGATACTACAATGCAATCAGTATCA
ACAAGTGGTTTGGCTTGGAAGATGGAGTCTAGGGGCTCTACAGGAGTAGCTACTCTCTAATG
GAGTTGCATTTTGAAGCAGGACACTGTGAAAAGCTGGCCTCCTAAAGAGGCTGCTAAACATT
AGGGTCAATTTTCCAGTGCACTTTCTGAAGTGTCTGCAGTTCCCCATGCAAAGCTGCCCAAA
CATAGCACTTCCAATTGAATACAATTATATGCAGGCGTACTGCTTCTTGCCAGCACTGTCCT
TCTCAAATGAACTCAACAAACAATTTCAAAGTCTAGTAGAAAGTAACAAGCTTTGAATGTCA
TTAAAAGTATATCTGCTTTCAGTAGTTCAGCTTATTTATGCCCACTAGAAACATCTTGTAC
AAGCTGAACACTGGGGCTCCAGATTAGTGGTAAAACCTACTTTATACAATCATAGAATCATA
GAATGGCCTGGGTTGGAAGGGACCCCAAGGATCATGAAGATCCAACACCCCCGCCACAGGCA
GGGCCACCAACCTCCAGATCTGGTACTAGACCAGGCAGCCCAGGGCTCCATCCAACCTGGCC
ATGAACACCTCCAGGGATGGAGCATCCACAACCTCTCTGGGCAGCCTGTGCCAGCACCTCAC
CACCCTCTCTGTGAAGAACTTTTCCCTGACATCCAATCTAAGCCTTCCCTCCTTGAGGTTAG
ATCCACTCCCCCTTGTGCTATCACTGTCTACTCTTGTAAAAGTTGATTCTCCTCCTTTTTG
GAAGGTTGCAATGAGGTCTCCTTGCAGCCTTCTTCTCTTCTGCAGGATGAACAAGCCCAGCT
CCCTCAGCCTGTCTTTATAGGAGAGGTGCTCCAGCCCTCTGATCATCTTTGTGGCCCTCCTC
TGGACCCGCTCCAAGAGCTCCACATCTTTCCTGTACTGGGGGCCCCAGGCCTGAATGCAGTA
CTCCAGATGGGGCCTCAAAAGAGCAGAGTAAAGAGGGACAATCACCTTCCTCACCCTGCTGG
CCAGCCCTCTTCTGATGGAGCCCTGGATACAACTGGCTTTCTGAGCTGCAACTTCTCCTTAT
CAGTTCCACTATTAAAACAGGAACAATACAACAGGTGCTGATGGCCAGTGCAGAGTTTTTCA
CACTTCTTCATTTCGGTAGATCTTAGATGAGGAACGTTGAAGTTGTGCTTCTGCGTGTGCTT
CTTCCTCCTCAAATACTCCTGCCTGATACCTCACCCCACCTGCCACTGAATGGCTCCATGGC
CCCCTGCAGCCAGGGCCCTGATGAACCCGGCACTGCTTCAGATGCTGTTTAATAGCACAGTA
TGACCAAGTTGCACCTATGAATACACAAACAATGTGTTGCATCCTTCAGCACTTGAGAAGAA
GAGCCAAATTTGCATTGTCAGGAAATGGTTTAGTAATTCTGCCAATTAAAACTTGTTTATCT
ACCATGGCTGTTTTTATGGCTGTTAGTAGTGGTACACTGATGATGAACAATGGCTATGCAGT
AAAATCAAGACTGTAGATATTGCAACAGACTATAAAATTCCTCTGTGGCTTAGCCAATGTGG
TACTTCCCACATTGTATAAGAAATTTGGCAAGTTTAGAGCAATGTTTGAAGTGTTGGGAAAT
TTCTGTATACTCAAGAGGGCGTTTTTGACAACTGTAGAACAGAGGAATCAAAAGGGGGTGGG
AGGAAGTTAAAAGAAGAGGCAGGTGCAAGAGAGCTTGCAGTCCCGCTGTGTGTACGACACTG
GCAACATGAGGTCTTTGCTAATCTTGGTGCTTTGCTTCCTGCCCCTGGCTGCCTTAGGGTGC
GATCTGCCTCAGACCCACAGCCTGGGCAGCAGGAGGACCCTGATGCTGCTGGCTCAGATGAG
GAGAATCAGCCTGTTTAGCTGCCTGAAGGATAGGCACGATTTTGGCTTTCCTCAAGAGGAGT
TTGGCAACCAGTTTCAGAAGGCTGAGACCATCCCTGTGCTGCACGAGATGATCCAGCAGATC
TTTAACCTGTTTAGCACCAAGGATAGCAGCGCTGCTTGGGATGAGACCCTGCTGGATAAGTT
TTACACCGAGCTGTACCAGCAGCTGAACGATCTGGAGGCTTGCGTGATCCAGGGCGTGGGCG
TGACCGAGACCCCTCTGATGAAGGAGGATAGCATCCTGGCTGTGAGGAAGTACTTTCAGAGG
ATCACCCTGTACCTGAAGGAGAAGAAGTACAGCCCCTGCGCTTGGGAAGTCGTGAGGGCTGA
GATCATGAGGAGCTTTAGCCTGAGCACCAACCTGCAAGAGAGCTTGAGGTCTAAGGAGTAAA
AAGTCTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTT
GGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTAT
TGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATT
TTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAA

```
TGTGGTAAAATCGATAAGGATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCT
CCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATG
CAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCT
CGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACA
GAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCG
TAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAA
ATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCC
CCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGC
CTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGG
TGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGC
GCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGC
AGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGA
AGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAG
CGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATC
CTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTG
GTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAA
ATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGG
CACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG
ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCC
ACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAA
GTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTA
AGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTC
ACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACAT
GATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGT
AAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCAT
GCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT
GTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGC
AGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTT
ACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTT
TTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGA
ATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCAT
TTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAA
TAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGCGCCCTGTAGCGGCGCATTA
AGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCC
CGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTC
TAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAA
CTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTT
GACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACC
CTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAA
AATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATTTC
CCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTAT
TACGCCAGCCCAAGCTACCATGATAAGTAAGTAATATTAAGGTACGGGAGGTACTTGGAGCG
GCCGCTCTAGAACTAGTGGATCCCCCGGCCGCAATAAAATATCTTTATTTTCATTACATCTG
TGTGTTGGTTTTTTGTGTGAATCGATAGTACTAACATACGCTCTCCATCAAAACAAAACGAA
ACAAACAAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCT
ATCGATAGGTACCGAGCTCTTACGCGTGCTAGCCCTCGAGCAGGATCTATACATTGAATCAA
TATTGGCAATTAGCCATATTAGTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCC
ATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCCAATATGAC
CGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTT
CATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
```

GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACAT
CAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTG
GCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAG
TCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTT
GACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA
AAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTA
GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGG
AGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCCCTC
GAAGCTCGACTCTAGGGGCTCGAGATCTGCGATCTAAGTAAGCTTGCATGCCTGCAGGTCGG
CCGCCACGACCGGTGCCGCCACCATCCCCTGACCCACGCCCCTGACCCCTCACAAGGAGACG
ACCTTCCATGACCGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGG
CCGTACGCACCCTCGCCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGACCCG
GACCGCCACATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGA
CATCGGCAAGGTGTGGGTCGCGGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGA
GCGTCGAAGCGGGGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCC
CGGCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGC
GTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCG
TCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCC
GCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGTCGAGGT
GCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCCCGCCCCACG
ACCCGCAGCGCCCGACCGAAAGGAGCGCACGACCCCATGGCTCCGACCGAAGCCGACCCGGG
CGGCCCCGCCGACCCCGCACCCGCCCCCGAGGCCCACCGACTCTAGAGTCGGGGCGGCCGGC
CGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAG
TGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAG
CTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGG
TGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCGATAAGGATCAA
TTCGGCTTCAGGTACCGTCGACGATGTAGGTCACGGTCTCGAAGCCGCGGTGCGGGTGCCAG
GGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACCTCACCCATCTGGTCCATCATGATGAA
CGGGTCGAGGTGGCGGTAGTTGATCCCGGCGAACGCGCGGCGCACCGGGAAGCCCTCGCCCT
CGAAACCGCTGGGCGCGGTGGTCACGGTGAGCACGGGACGTGCGACGGCGTCGGCGGGTGCG
GATACGCGGGGCAGCGTCAGCGGGTTCTCGACGGTCACGGCGGGCATGTCGACAGCCGAATT
GATCCGTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGG
GCATGACTATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTG
CCGGCAGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGC
GAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCA
GGAAAGAACATG

Fig. 15

pOM IFN-Ins-CMV-pur-attB (SEQ ID NO: 8)

```
GGCCGCCACCGCGGTGGAGCTCCAATTCGCCCTATAGTGAGTCGTATTACAATTCACTGGCC
GTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGC
ACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAAC
AGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGT
GTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGC
TTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGC
TCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGT
GATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTC
CACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCT
ATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATT
TAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGCTTACAATTTAGGTGGCACTTTT
CGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAATATGTATCC
GCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGTA
TTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCT
CACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTTA
CATCGAACTGGATCTCAACAGCGGTAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTC
CAATGATGAGCACTTTTAAAGTTCTGCTATGTGGCGCGGTATTATCCCGTATTGACGCCGGG
CAAGAGCAACTCGGTCGCCGCATACACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGT
CACAGAAAAGCATCTTACGGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCA
TGAGTGATAACACTGCGGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACC
GCTTTTTTGCACAACATGGGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAA
TGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGC
GCAAACTATTAACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATG
GAGGCGGATAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGC
TGATAAATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATG
GTAAGCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGA
AATAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGT
TTACTCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGA
AGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCG
TCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTG
CTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTAC
CAACTCTTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTA
GTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCT
GCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACT
CAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAG
CCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAG
CGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAG
GAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTT
CGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAA
AAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTGCTGGCCTTTTGCTCACATGT
TCTTTCCTGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGAT
ACCGCTCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAGAGCG
CCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACA
GGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAT
TAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGG
ATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTCGAAATTAACCCTC
ACTAAAGGGAACAAAAGCTGGGTACCGGGCCCCCCCTCGACTAGAGGGACAGCCCCCCCCA
AAGCCCCCAGGGATGTAATTACGTCCCTCCCCGCTAGGGGGCAGCAGCGAGCCGCCCGGGG
CTCCGCTCCGGTCCGGCGCTCCCCCCGCATCCCCGAGCCGGCAGCGTGCGGGACAGCCCGG
GCACGGGGAAGGTGGCACGGGATCGCTTTCCTCTGAACGCTTCTCGCTGCTCTTTGAGCCTG
```

```
CAGACACCTGGGGGGGATACGGGGAAAAAGCTTTAGGCTGAAAGAGAGATTTAGAATGACAGA
ATCATAGAACGGCCTGGGTTGCAAAGGAGCACAGTGCTCATCCAGATCCAACCCCCTGCTAT
GTGCAGGGTCATCAACCAGCAGCCCAGGCTGCCCAGAGCCACATCCAGCCTGGCCTTGAATG
CCTGCAGGGATGGGGCATCCACAGCCTCCTTGGGCAACCTGTTCAGTGCGTCACCACCCTCT
GGGGGAAAAACTGCCTCCTCATATCCAACCCAAACCTCCCCTGTCTCAGTGTAAAGCCATTC
CCCCTTGTCCTATCAAGGGGGAGTTTGCTGTGACATTGTTGGTCTGGGGTGACACATGTTTG
CCAATTCAGTGCATCACGGAGAGGCAGATCTTGGGGATAAGGAAGTGCAGGACAGCATGGAC
GTGGGACATGCAGGTGTTGAGGGCTCTGGGACACTCTCCAAGTCACAGCGTTCAGAACAGCC
TTAAGGATAAGAAGATAGGATAGAAGGACAAAGAGCAAGTTAAAACCCAGCATGGAGAGGAG
CACAAAAAGGCCACAGACACTGCTGGTCCCTGTGTCTGAGCCTGCATGTTTGATGGTGTCTG
GATGCAAGCAGAAGGGGTGGAAGAGCTTGCCTGGAGAGATACAGCTGGGTCAGTAGGACTGG
GACAGGCAGCTGGAGAATTGCCATGTAGATGTTCATACAATCGTCAAATCATGAAGGCTGGA
AAAGCCCTCCAAGATCCCCAAGACCAACCCCAACCCACCCACCGTGCCCACTGGCCATGTCC
CTCAGTGCCACATCCCCACAGTTCTTCATCACCTCCAGGGACGGTGACCCCCCCACCTCCGT
GGGCAGCTGTGCCACTGCAGCACCGCTCTTTGGAGAAGGTAAATCTTGCTAAATCCAGCCCG
ACCCTCCCCTGGCACAACGTAAGGCCATTATCTCTCATCCAACTCCAGGACGGAGTCAGTGA
GGATGGGGCTCTAGTCGAGGTCGACGGTATCGATAAGCTTGATTAGGCAGAGCAATAGGACT
CTCAACCTCGTGAGTATGGCAGCATGTTAACTCTGCACTGGAGTCCAGCGTGGGAAACAATC
TGCCTTGCACATGAGTCTTCGTGGGCCAATATTCCCCAACGGTTTTCCTTCAGCTTGTCTTG
TCTCCTAAGCTCTCAAAACACCTTTTTGGTGAATAAACTCACTTGGCAACGTTTATCTGTCT
TACCTTAGTGTCACGTTTCATCCCTATTCCCCTTTCTCCTCCTCCGTGTGGTACACAGTGGT
GCACACTGGTTCTTCTGTTGATGTTCTGCTCTGACAGCCAATGTGGGTAAAGTTCTTCCTGC
CACGTGTCTGTGTTGTTTTCACTTCAAAAAGGGCCCTGGGCTCCCCTTGGAGCTCTCAGGCA
TTTCCTTAATCATCACAGTCACGCTGGCAGGATTAGTCCCTCCTAAACCTTAGAATGACCTG
AACGTGTGCTCCCTCTTTGTAGTCAGTGCAGGGAGACGTTTGCCTCAAGATCAGGGTCCATC
TCACCCACAGGGCCATTCCCAAGATGAGGTGGATGGTTTACTCTCACAAAAAGTTTTCTTAT
GTTTGGCTAGAAAGGAGAACTCACTGCCTACCTGTGAATTCCCCTAGTCCTGGTTCTGCTGC
CACTGCTGCCTGTGCAGCCTGTCCCATGGAGGGGGCAGCAACTGCTGTCACAAAGGTGATCC
CACCCTGTCTCCACTGAAATGACCTCAGTGCCACGTGTTGTATAGGGTATAAAGTACGGGAG
GGGGATGCCCGGCTCCCTTCAGGGTTGCAGAGCAGAAGTGTCTGTGTATAGAGTGTGTCTTA
ATCTATTAATGTAACAGAACAACTTCAGTCCTAGTGTTTTGTGGGCTGGAATTGCCCATGTG
GTAGGGACAGGCCTGCTAAATCACTGCAATCGCCTATGTTCTGAAGGTATTTGGGAAAGAAA
GGGATTTGGGGGATTGCCTGTGATTGGCTTTAATTGAATGGCAAATCACAGGAAAGCAGTTC
TGCTCAACAGTTGGTTGTTTCAGCCAATTCTTGCAGCCAAAGAGCCGGGTGCCCAGCGATAT
AATAGTTGTCACTTGTGTCTGTATGGATGACAGGGAGGTAGGGTGACCTGAGGACCACCCTC
CAGCTTCTGCTAGCGTAGGTACAGTCACCACCTCCAGCTCCACACGAGTCCCATCGTGGTTT
ACCAAAGAAACACAATTATTTGGACCAGTTTGGAAAGTCACCCGCTGAATTGTGAGGCTAGA
TTAATAGAGCTGAAGAGCAAATGTTCCCAACTTGGAGATACTAGTTGGTATTAGTATCAGAG
GAACAGGGCCATAGCACCTCCATGCTATTAGATTCCGGCTGGCATGTACTTTTCAAGATGAT
TTGTAACTAACAATGGCTTATTGTGCTTGTCTTAAGTCTGTGTCCTAATGTAAATGTTCCTT
TGGTTTATATAACCTTCTTGCCATTTGCTCTTCAGGTGTTCTTGCAGAACACTGGCTGCTTT
AATCTAGTTTAACTGTTGCTTGATTATTCTTAGGGATAAGATCTGAATAAACTTTTTGTGGC
TTTGGCAGACTTTAGCTTGGGCTTAGCTCCCACATTAGCTTTTGCTGCCTTTTCTGTGAAGC
TATCAAGATCCTACTCAATGACATTAGCTGGGTGCAGGTGTACCAAATCCTGCTCTGTGGAA
CACATTGTCTGATGATACCGAAGGCAAACGTGAACTCAAAGAGGCACAGAGTTAAGAAGAAG
TCTGTGCAATTCAGAGGAAAAGCCAAAGTGGCCATTAGACACACTTTCCATGCAGCATTTGC
CAGTAGGTTTCATATAAAACTACAAAATGGAATAAACCACTACAAATGGGAAAAGCCTGATA
CTAGAATTTAAATATTCACCCAGGCTCAAGGGGTGTTTCATGGAGTAATATCACTCTATAAA
AGTAGGGCAGCCAATTATTCACAGACAAAGCTTTTTTTTTTCTGTGCTGCAGTGCTGTTTTT
CGGCTGATCCAGGGTTACTTATTGTGGGTCTGAGAGCTGAATGATTTCTCCTTGTGTCATGT
TGGTGAAGGAGATATGGCCAGGGGGAGATGAGCATGTTCAAGAGGAAACGTTGCATTTTGGT
GGCTTGGGAGAAAGGTAGAACGATATCAGGTCCATAGTGTCACTAAGAGATCTGAAGGATGG
```

```
TTTTACAGAACAGTTGACTTGGCTGGGTGCAGGCTTGGCTGTAAATGGATGGAAGGATGGAC
AGATGGGTGGACAGAGATTTCTGTGCAGGAGATCATCTCCTGAGCTCGGTGCTTGACAGACT
GCAGATCCATCCCATAACCTTCTCCAGCATGAGAGCGCGGGGAGCTTTGGTACTGTTCAGTC
TGCTGCTTGTTGCTTCCTGGGTGCACAGTGGTGATTTTCTTACTCACACAGGGCAAAAACCT
GAGCAGCTTCAAAGTGAACAGGTTGCTCTCATAGGCCATTCAGTTGTCAAGATGAGGTTTTT
GGTTTCTTGTTTTGTAAGGTGGGAAGAAGCACTGAAGGATCAGTTGCGAGGGCAGGGGTTTA
GCACTGTTCAGAGAAGTCTTATTTTAACTCCTCTCATGAACAAAAAGAGATGCAGGTGCAGA
TTCTGGCAAGCATGCAGTGAAGGAGAAAGCCCTGAATTTCTGATATATGTGCAATGTTGGGC
ACCTAACATTCCCCGCTGAAGCACAGCAGCTCCAGCTCCATGCAGTACTCACAGCTGGTGCA
GCCCTCGGCTCCAGGGTCTGAGCAGTGCTGGGACTCACGAGGTTCCATGTCTTTCACACTGA
TAATGGTCCAATTTCTGGAATGGGTGCCCATCCTTGGAGGTCCCCAAGGCCAGGCTGGCTGC
GTCTCCGAGCAGCCCGATCTGGTGGTGAGTAGCCAGCCCATGGCAGGAGTTAGAGCCTGATG
GTCTTTAAGGTCCCTTCCAACCTAAGCCATCCTACGATTCTAGGAATCATGACTTGTGAGTG
TGTATTGCAGAGGCAATATTTTAAAGTTATAAATGTTTTCTCCCCTTCCTTGTTTGTCAAAG
TTATCTTGATCGCCTTATCAATGCTTTTGGAGTCTCCAGTCATTTTTCTTACAMCAAAAAGA
GGAGGAAGAATGAAGAGAATCATTTAATTTCTTGATTGAATAGTAGGATTCAGAAAGCTGTA
CGTAATGCCGTCTCTTTGTATCGAGCTGTAAGGTTTCTCATCATTTATCAGCGTGGTACATA
TCAGCACTTTTCCATCTGATGTGGAAAAAAAATCCTTATCATCTACAGTCTCTGTACCTAA
ACATCGCTCAGACTCTTTACCAAAAAAGCTATAGGTTTTAAAACTACATCTGCTGATAATTT
GCCTTGTTTTAGCTCTTCTTCCATATGCTGCGTTTGTGAGAGGTGCGTGGATGGGCCTAAAC
TCTCAGCTGCTGAGCTTGATGGGTGCTTAAGAATGAAGCACTCACTGCTGAAACTGTTTTCA
TTTCACAGGAATGTTTTAGTGGCATTGTTTTTATAACTACATATTCCTCAGATAAATGAAAT
CCAGAAATAATTATGCAAACTCACTGCATCCGTTGCACAGGTCTTTATCTGCTAGCAAAGGA
AATAATTTGGGGATGGCAAAAACATTCCTTCAGACATCTATATTTAAAGGAATATAATCCTG
GTACCCACCCACTTCATCCCTCATTATGTTCACACTCAGAGATACTCATTCTCTTGTTGTTA
TCATTTGATAGCGTTTTCTTTGGTTCTTTGCCACGCTCTGGGCTATGGCTGCACGCTCTGCA
CTGATCAGCAAGTAGATGCGAGGGAAGCAGCAGTGAGAGGGGCTGCCCTCAGCTGGCACCCA
GCCGCTCAGCCTAGGAGGGGACCTTGCCTTTCCACCAGCTGAGGTGCAGCCCTACAAGCTTA
CACGTGCTGCGAGCAGGTGAGCAAAGGGAGTCTTCATGGTGTGTTTCTTGCTGCCCGGAAGC
AAAACTTTACTTTCATTCATTCCCCTTGAAGAATGAGGAATGTTTGGAAACGGACTGCTTTA
CGTTCAATTTCTCTCTTCCCTTTAAGGCTCAGCCAGGGGCCATTGCTGAGGACGGCATCGGG
GCCCCCTGGACCAAATCTGTGGCACAGATGGTTTCACTTACATCAGTGGATGTGGGATCTGC
GCCTGTAATGTGTCCTTCTGAAGGAAGGAACGTGCCTTCCAAGTGCCAGCCCCACAGCCCCC
AGCCCCTCCCTGTGCTGCTCCAATTCATCTCCTCTTCCTCCTTCTCCCTTTGCTGTTTGTGC
TCGGGTAGAAATCATGAAGATTTAGAAGAGAAAACAAAATAACTGGAGTGGAAACCCAGGTG
ATGCAGTTCATTCAGCTGTCATAGGTTTGTCGTTGCTATAGGTCTGTATCAGAGATGCTARC
ACCACTTTGCTGTCGGTGCTTAACTCGGGTGAACTCTCCTTCACTCGCATCATTTGCGGGCC
TTATTTACATCCCCAGCATCCATCACCCTCTGGGAAAATGGGCGCACTGGATCTCTAATGGA
AGACTTTCCCTCTTTCAGAGCCTGTGGGATGTGCAGTGACAAGAAACGTGGAGGGGCTGAGC
AGCAGCACTGCCCCCAGGGAGCAGGAGCGGATGCCATCGGTGGCAGCATCCCAAATGATGTC
AGCGGATGCTGAGCAGGCAGCGGACGAACGGACAGAAGCGATGCGTACACCTTCTGTTGACA
TGGTATTTGGCAGCGATTTAACACTCGCTTCCTAGTCCTGCTATTCTCCACAGGCTGCATTC
AAATGAACGAAGGGAAGGGAGGCAAAAAGATGCAAAATCCGAGACAAGCAGCAGAAATATTT
CTTCGCTACGGAAGCGTGCGCAAACAACCTTCTCCAACAGCACCAGAAGAGCACAGCGTAAC
CTTTTTCAAGACCAGAAAAGGAAATTCACAAAGCCTCTGTGGATACCAGCGCGTTCAGCTCT
CCTGATAGCAGATTTCTTGTCAGGTTGCGAATGGGGTATGGTGCCAGGAGGTGCAGGGACCA
TATGATCATATACAGCACAGCAGTCATTGTGCATGTATTAATATATATTGAGTAGCAGTGTT
ACTTTGCCAAAGCAATAGTTCAGAGATGAGTCCTGCTGCATACCTCTATCTTAAAACTAACT
TATAAATAGTAAAACCTTCTCAGTTCAGCCACGTGCTCCTCTCTGTCAGCACCAATGGTGCT
TCGCCTGCACCCAGCTGCAAGGAATCAGCCCGTGATCTCATTAACACTCAGCTCTGCAGGAT
AAATTAGATTGTTCCACTCTCTTTTGTTGTTAATTACGACGGAACAATTGTTCAGTGCTGAT
GGTCCTAATTGTCAGCTACAGAAAACGTCTCCATGCAGTTCCTTCTGCGCCAGCAAACTGTC
```

```
CAGGCTATAGCACCGTGATGCATGCTACCTCTCACTCCATCCTTCTTCTCTTTCCCACCAGG
GAGAGCTGTGTGTTTTCACTCTCAGCCACTCTGAACAATACCAAACTGCTACGCACTGCCTC
CCTCGGAAAGAGAATCCCCTTGTTGCTTTTTTATTTACAGGATCCTTCTTAAAAAGCAGACC
ATCATTCACTGCAAACCCAGAGCTTCATGCCTCTCCTTCCACAACCGAAAACAGCCGGCTTC
ATTTGTCTTTTTTAAATGCTGTTTTCCAGGTGAATTTTGGCCAGCGTGTTGGCTGAGATCCA
GGAGCACGTGTCAGCTTTCTGCTCTCATTGCTCCTGTTCTGCATTGCCTCTTTCTGGGGTTT
CCAAGAGGGGGGGAGACTTTGCGCGGGGATGAGATAATGCCCCTTTTCTTAGGGTGGCTGCT
GGGCAGCAGAGTGGCTCTGGGTCACTGTGGCACCAATGGGAGGCACCAGTGGGGGTGTGTTT
TGTGCAGGGGGGAAGCATTCACAGAATGGGGCTGATCCTGAAGCTTGCAGTCCAAGGCTTTG
TCTGTGTACCCAGTGAAATCCTTCCTCTGTTACATAAAGCCCAGATAGGACTCAGAAATGTA
GTCATTCCAGCCCCCTCTTCCTCAGATCTGGAGCAGCACTTGTTTGCAGCCAGTCCTCCCC
AAAATGCACAGACCTCGCCGAGTGGAGGGAGATGTAAACAGCGAAGGTTAATTACCTCCTTG
TCAAAAACACTTTGTGGTCCATAGATGTTTCTGTCAATCTTACAAAACAGAACCGAGAGGCA
GCGAGCACTGAAGAGCGTGTTCCCATGCTGAGTTAATGAGACTTGGCAGCTCGCTGTGCAGA
GATGATCCCTGTGCTTCATGGGAGGCTGTAACCTGTCTCCCCATCGCCTTCACACCGCAGTG
CTGTCCTGGACACCTCACCCTCCATAAGCTGTAGGATGCAGCTGCCCAGGGATCAAGAGACT
TTTCCTAAGGCTCTTAGGACTCATCTTTGCCGCTCAGTAGCGTGCAGCAATTACTCATCCCA
ACTATACTGAATGGGTTTCTGCCAGCTCTGCTTGTTTGTCAATAAGCATTTCTTCATTTTGC
CTCTAAGTTTCTCTCAGCAGCACCGCTCTGGGTGACCTGAGTGGCCACCTGGAACCCGAGGG
GCACAGCCACCACCTCCCTGTTGCTGCTGCTCCAGGGACTCATGTGCTGCTGGATGGGGGGA
AGCATGAAGTTCCTCACCCAGACACCTGGGTTGCAATGGCTGCAGCGTGCTCTTCTTGGTAT
GCAGATTGTTTCCAGCCATTACTTGTAGAAATGTGCTGTGGAAGCCCTTTGTATCTCTTTCT
GTGGCCCTTCAGCAAAAGCTGTGGGAAAGCTCTGAGGCTGCTTTCTTGGGTCGTGGAGGAAT
TGTATGTTCCTTCTTTAACAAAAATTATCCTTAGGAGAGAGCACTGTGCAAGCATTGTGCAC
ATAAAACAATTCAGGTTGAAAGGGCTCTCTGGAGGTTTCCAGCCTGACTACTGCTCGAAGCA
AGGCCAGGTTCAAAGATGGCTCAGGATGCTGTGTGCCTTCCTGATTATCTGTGCCACCAATG
GAGGAGATTCACAGCCACTCTGCTTCCCGTGCCACTCATGGAGAGGAATATTCCCTTATATT
CAGATAGAATGTTATCCTTTAGCTCAGCCTTCCCTATAACCCCATGAGGGAGCTGCAGATCC
CCATACTCTCCCCTTCTCTGGGGTGAAGGCCGTGTCCCCCAGCCCCCCTTCCCACCCTGTGC
CCTAAGCAGCCCGCTGGCCTCTGCTGGATGTGTGCCTATATGTCAATGCCTGTCCTTGCAGT
CCAGCCTGGGACATTTAATTCATCACCAGGGTAATGTGGAACTGTGTCATCTTCCCCTGCAG
GGTACAAAGTTCTGCACGGGGTCCTTTCGGTTCAGGAAAACCTTCACTGGTGCTACCTGAAT
CAAGCTCTATTTAATAAGTTCATAAGCACATGGATGTGTTTTCCTAGAGATACGTTTTAATG
GTATCAGTGATTTTTATTTGCTTTGTTGCTTACTTCAAACAGTGCCTTTGGGCAGGAGGTGA
GGGACGGGTCTGCCGTTGGCTCTGCAGTGATTTCTCCAGGCGTGTGGCTCAGGTCAGATAGT
GGTCACTCTGTGGCCAGAAGAAGGACAAAGATGGAAATTGCAGATTGAGTCACGTTAAGCAG
GCATCTTGGAGTGATTTGAGGCAGTTTCATGAAAGAGCTACGACCACTTATTGTTGTTTTCC
CCTTTTACAACAGAAGTTTTCATCAAAATAACGTGGCAAAGCCCAGGAATGTTTGGGAAAAG
TGTAGTTAAATGTTTTGTAATTCATTTGTCGGAGTGCTACCAGCTAAGAAAAAGTCCTACC
TTTGGTATGGTAGTCCTGCAGAGAATACAACATCAATATTAGTTTGGAAAAAAACACCACCA
CCACCAGAAACTGTAATGGAAAATGTAAACCAAGAAATTCCTTGGGTAAGAGAGAAAGGATG
TCGTATACTGGCCAAGTCCTGCCCAGCTGTCAGCCTGCTGACCCTCTGCAGTTCAGGACCAT
GAAACGTGGCACTGTAAGACGTGTCCCCTGCCTTTGCTTGCCCACAGATCTCTGCCCTTGTG
CTGACTCCTGCACACAAGAGCATTTCCCTGTAGCCAAACAGCGATTAGCCATAAGCTGCACC
TGACTTTGAGGATTAAGAGTTTGCAATTAAGTGGATTGCAGCAGGAGATCAGTGGCAGGGTT
GCAGATGAAATCCTTTTCTAGGGGTAGCTAAGGGCTGAGCAACCTGTCCTACAGCACAAGCC
AAACCAGCCAAGGGTTTTCCTGTGCTGTTCACAGAGGCAGGGCCAGCTGGAGCTGGAGGAGG
TTGTGCTGGGACCCTTCTCCCTGTGCTGAGAATGGAGTGATTTCTGGGTGCTGTTCCTGTGG
CTTGCACTGAGCAGCTCAAGGGAGATCGGTGCTCCTCATGCAGTGCCAAAACTCGTGTTTGA
TGCAGAAAGATGGATGTGCACCTCCCTCCTGCTAATGCAGCCGTGAGCTTATGAAGGCAATG
AGCCCTCAGTGCAGCAGGAGCTGTAGTGCACTCCTGTAGGTGCTAGGGAAAATCTCTGGTTC
CCAGGGATGCATTCATAAGGGCAATATATCTTGAGGCTGCGCCAAATCTTTCTGAAATATTC
```

ATGCGTGTTCCCTTAATTTATAGAAACAAACACAGCAGAATAATTATTCCAATGCCTCCCCT
CGAAGGAAACCCATATTTCCATGTAGAAATGTAACCTATATACACACAGCCATGCTGCATCC
TTCAGAACGTGCCAGTGCTCATCTCCCATGGCAAAATACTACAGGTATTCTCACTATGTTGG
ACCTGTGAAAGGAACCATGGTAAGAAACTTCGGTTAAAGGTATGGCTGCAAAACTACTCATA
CCAAAACAGCAGAGCTCCAGACCTCCTCTTAGGAAAGAGCCACTTGGAGAGGGATGGTGTGA
AGGCTGGAGGTGAGAGACAGAGCCTGTCCCAGTTTTCCTGTCTCTATTTTCTGAAACGTTTG
CAGGAGGAAAGGACAACTGTACTTTCAGGCATAGCTGGTGCCCTCACGTAAATAAGTTCCCC
GAACTTCTGTGTCATTTGTTCTTAAGATGCTTTGGCAGAACACTTTGAGTCAATTCGCTTAA
CTGTGACTAGGTCTGTAAATAAGTGCTCCCTGCTGATAAGGTTCAAGTGACATTTTTAGTGG
TATTTGACAGCATTTACCTTGCTTTCAAGTCTTCTACCAAGCTCTTCTATACTTAAGCAGTG
AAACCGCCAAGAAACCCTTCCTTTTATCAAGCTAGTGCTAAATACCATTAACTTCATAGGTT
AGATACGGTGCTGCCAGCTTCACCTGGCAGTGGTTGGTCAGTTCTGCTGGTGACAAAGCCTC
CCTGGCCTGTGCTTTTACCTAGAGGTGAATATCCAAGAATGCAGAACTGCATGGAAAGCAGA
GCTGCAGGCACGATGGTGCTGAGCCTTAGCTGCTTCCTGCTGGGAGATGTGGATGCAGAGAC
GAATGAAGGACCTGTCCCTTACTCCCCTCAGCATTCTGTGCTATTTAGGGTTCTACCAGAGT
CCTTAAGAGGTTTTTTTTTTTTTGGTCCAAAAGTCTGTTTGTTTGGTTTTGACCACTGAGA
GCATGTGACACTTGTCTCAAGCTATTAACCAAGTGTCCAGCCAAAATCAATTGCCTGGGAGA
CGCAGACCATTACCTGGAGGTCAGGACCTCAATAAATATTACCAGCCTCATTGTGCCGCTGA
CAGATTCAGCTGGCTGCTCCGTGTTCCAGTCCAACAGTTCGGACGCCACGTTTGTATATATT
TGCAGGCAGCCTCGGGGGGACCATCTCAGGAGCAGAGCACCGGCAGCCGCCTGCAGAGCCGG
GCAGTACCTCACCATGGCTTTGACCTTTGCCTTACTGGTGGCTCTCCTGGTGCTGAGCTGCA
AGAGCAGCTGCTCTGTGGGCTGCGATCTGCCTCAGACCCACAGCCTGGGCAGCAGGAGGACC
CTGATGCTGCTGGCTCAGATGAGGAGAATCAGCCTGTTTAGCTGCCTGAAGGATAGGCACGA
TTTTGGCTTTCCTCAAGAGGAGTTTGGCAACCAGTTTCAGAAGGCTGAGACCATCCCTGTGC
TGCACGAGATGATCCAGCAGATCTTTAACCTGTTTAGCACCAAGGATAGCAGCGCTGCTTGG
GATGAGACCCTGCTGGATAAGTTTTACACCGAGCTGTACCAGCAGCTGAACGATCTGGAGGC
TTGCGTGATCCAGGGCGTGGGCGTGACCGAGACCCCTCTGATGAAGGAGGATAGCATCCTGG
CTGTGAGGAAGTACTTTCAGAGGATCACCCTGTACCTGAAGGAGAAGAAGTACAGCCCCTGC
GCTTGGGAAGTCGTGAGGGCTGAGATCATGAGGAGCTTTAGCCTGAGCACCAACCTGCAAGA
GAGCTTGAGGTCTAAGGAGTAAAAAGTCTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACA
TGATAAGATACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTT
ATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGT
TAACAACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTT
AAAGCAAGTAAAACCTCTACAAATGTGGTAAAATCGATACCGTCGACCTCGACTAGAGCGGC
CACTAACATACGCTCTCCATCAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTC
CCCAGTGCAAGTGCAGGTGCCAGAACATTTCTCTATCGATAGGTACCGAGCTCTTACGCGTG
CTAGCCCTCGAGCAGGATCTATACATTGAATCAATATTGGCAATTAGCCATATTAGTCATTG
GTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAAT
ATGTACATTTATATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGACTAG
TTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTA
CATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCA
ATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGA
GTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCC
CTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGG
GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTT
TTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACC
CCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT
AACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG
CAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCC
ATAGAAGACACCGGGACCGATCCAGCCTCCCCTCGAAGCTCGACTCTAGGGGCTCGAGATCT
GCGATCTAAGTAAGCTTGCATGCCTGCAGGTCGGCCGCCACGACCGGTGCCGCCACCATCCC
CTGACCCACGCCCCTGACCCCTCACAAGGAGACGACCTTCCATGACCGAGTACAAGCCCACG

```
GTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCACCCTCGCCGCCGCGTTCGC
CGACTACCCCGCCACGCGCCACACCGTCGACCCGGACCGCCACATCGAGCGGGTCACCGAGC
TGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGGCAAGGTGTGGGTCGCGGACGAC
GGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGTCGAAGCGGGGGCGGTGTTCGCCGA
GATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCGGCTGGCCGCGCAGCAACAGATGGAAG
GCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGCGTGGTTCCTGGCCACCGTCGGCGTCTCG
CCCGACCACCAGGGCAAGGGTCTGGGCAGCGCCGTCGTGCTCCCCGGAGTGGAGGCGGCCGA
GCGCGCCGGGGTGCCCGCCTTCCTGGAGACCTCCGCGCCCCGCAACCTCCCCTTCTACGAGC
GGCTCGGCTTCACCGTCACCGCCGACGTCGAGGTGCCCGAAGGACCGCGCACCTGGTGCATG
ACCCGCAAGCCCGGTGCCTGACGCCCGCCCCACGACCCGCAGCGCCCGACCGAAAGGAGCGC
ACGACCCCATGGCTCCGACCGAAGCCGACCCGGGCGGCCCCGCCGACCCCGCACCCGCCCCC
GAGGCCCACCGACTCTAGAGTCGGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACAT
TGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTT
GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAAT
TGCATTCATTTTATGTTTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAA
CCTCTACAAATGTGGTAAAATCGATAAGGATCAATTCGGCTTCAGGTACCGTCGACGATGTA
GGTCACGGTCTCGAAGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGT
ACTCCACCTCACCCATCTGGTCCATCATGATGAACGGGTCGAGGTGGCGGTAGTTGATCCCG
GCGAACGCGCGGCGCACCGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGT
GAGCACGGGACGTGCGACGGCGTCGGCGGGTGCGGATACGCGGGGCAGCGTCAGCGGGTTCT
CGACGGTCACGGCGGGCATGTCGACAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCC
TTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGAC
TGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGC
```

Fig. 16

pRSV-C31int (SEQ ID NO: 9)

```
CTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCC
GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT
CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG
TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC
CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA
AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT
CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG
GCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG
CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT
CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC
AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC
TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC
GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT
TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC
TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG
AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA
ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA
CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG
ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC
AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT
AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATC
GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG
CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC
GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT
TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG
TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT
AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG
CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA
AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA
TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG
CCACCTGACGTCGACGGATCGGGAGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAA
TCTGCTCTGATGCCGCATAGTTAAGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCG
CTGAGTAGTGCGCGAGCAAAATTTAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCA
TGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATA
CGCGTGCTAGGGGTCTAGGATCGATTCTAGGAATTCTCTAGCCGCGGTCTAGGGATCCCG
GCGCGTATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGTATCT
GCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATTTAAGCTACAAC
AAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGGCGTTTTGCGCT
GCTTCGCGATGTACGGGCCAGATATACGCGTATCTGAGGGGACTAGGGTGTGTTTAGGCG
AAAAGCGGGGCTTCGGTTGTACGCGGTTAGGAGTCCCCTCAGGATATAGTAGTTTCGCTT
TTGCATAGGGAGGGGGAAATGTAGTCTTATGCAATACACTTGTAGTCTTGCAACATGGTA
ACGATGAGTTAGCAACATGCCTTACAAGGAGAGAAAAGCACCGTGCATGCCGATTGGTG
GAAGTAAGGTGGTACGATCGTGCCTTATTAGGAAGGCAACAGACAGGTCTGACATGGATT
GGACGAACCACTGAATTCCGCATTGCAGAGATAATTGTATTTAAGTGCCTAGCTCGATAC
AATAAACGCCATTTGACCATTCACCACATTGGTGTGCACCTCCAAGCTTGCATGCCTGCA
GGTACCGGTCCGGAATTCCCGGGTCGACGAGCTCACTAGTCGTAGGGTCGCCGACATGAC
ACAAGGGGTTGTGACCGGGGTGGACACGTACGCGGGTGCTTACGACCGTCAGTCGCGCGA
```

GCGCGAGAATTCGAGCGCAGCAAGCCCAGCGACACAGCGTAGCGCCAACGAAGACAAGGC
GGCCGACCTTCAGCGCGAAGTCGAGCGCGACGGGGGCCGGTTCAGGTTCGTCGGGCATTT
CAGCGAAGCGCCGGGCACGTCGGCGTTCGGGACGGCGGAGCGCCCGGAGTTCGAACGCAT
CCTGAACGAATGCCGCGCCGGGCGGCTCAACATGATCATTGTCTATGACGTGTCGCGCTT
CTCGCGCCTGAAGGTCATGGACGCGATTCCGATTGTCTCGGAATTGCTCGCCCTGGGCGT
GACGATTGTTTCCACTCAGGAAGGCGTCTTCCGGCAGGGAAACGTCATGGACCTGATTCA
CCTGATTATGCGGCTCGACGCGTCGCACAAAGAATCTTCGCTGAAGTCGGCGAAGATTCT
CGACACGAAGAACCTTCAGCGCGAATTGGGCGGGTACGTCGGCGGGAAGGCGCCTTACGG
CTTCGAGCTTGTTTCGGAGACGAAGGAGATCACGCGCAACGGCCGAATGGTCAATGTCGT
CATCAACAAGCTTGCGCACTCGACCACTCCCCTTACCGGACCCTTCGAGTTCGAGCCCGA
CGTAATCCGGTGGTGGTGGCGTGAGATCAAGACGCACAAACACCTTCCCTTCAAGCCGGG
CAGTCAAGCCGCCATTCACCCGGGCAGCATCACGGGGCTTTGTAAGCGCATGGACGCTGA
CGCCGTGCCGACCCGGGGCGAGACGATTGGGAAGAAGACCGCTTCAAGCGCCTGGGACCC
GGCAACCGTTATGCGAATCCTTCGGGACCCGCGTATTGCGGGCTTCGCCGCTGAGGTGAT
CTACAAGAAGAAGCCGGACGGCACGCCGACCACGAAGATTGAGGGTTACCGCATTCAGCG
CGACCCGATCACGCTCCGGCCGGTCGAGCTTGATTGCGGACCGATCATCGAGCCCGCTGA
GTGGTATGAGCTTCAGGCGTGGTTGGACGGCAGGGGGCGCGGCAAGGGGCTTTCCCGGGG
GCAAGCCATTCTGTCCGCCATGGACAAGCTGTACTGCGAGTGTGGCGCCGTCATGACTTC
GAAGCGCGGGGAAGAATCGATCAAGGACTCTTACCGCTGCCGTCGCCGGAAGGTGGTCGA
CCCGTCCGCACCTGGGCAGCACGAAGGCACGTGCAACGTCAGCATGGCGGCACTCGACAA
GTTCGTTGCGGAACGCATCTTCAACAAGATCAGGCACGCCGAAGGCGACGAAGAGACGTT
GGCGCTTCTGTGGGAAGCCGCCCGACGCTTCGGCAAGCTCACTGAGGCGCCTGAGAAGAG
CGGCGAACGGGCGAACCTTGTTGCGGAGCGCGCCGACGCCCTGAACGCCCTTGAAGAGCT
GTACGAAGACCGCGCGGCAGGCGCGTACGACGGACCCGTTGGCAGGAAGCACTTCCGGAA
GCAACAGGCAGCGCTGACGCTCCGGCAGCAAGGGGCGGAAGAGCGGCTTGCCGAACTTGA
AGCCGCCGAAGCCCCGAAGCTTCCCCTTGACCAATGGTTCCCCGAAGACGCCGACGCTGA
CCCGACCGGCCCTAAGTCGTGGTGGGGGCGCGCGTCAGTAGACGACAAGCGCGTGTTCGT
CGGGCTCTTCGTAGACAAGATCGTTGTCACGAAGTCGACTACGGGCAGGGGGCAGGGAAC
GCCCATCGAGAAGCGCGCTTCGATCACGTGGGCGAAGCCGCCGACCGACGACGACGAAGA
CGACGCCCAGGACGGCACGGAAGACGTAGCGGCGTAGCGAGACACCCGGATCCCTCGAGG
GGCCCTATTCTATAGTGTCACCTAAATGCTAGAGCTCGCTGATCAGCCTCGACTGTGCCT
TCTAGTTGCCAGCCATCTGTTGTTTGCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGT
GCCACTCCCACTGTCCTTTCCTAATAAAATGAGGAAATTGCATCGCATTGTCTGAGTAGG
TGTCATTCTATTCTGGGGGGTGGGGTGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGAC
AATAGCAGGCATGCTGGGGATGCGGTGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGG
TGCCCAGTCATAGCCGAATAGCCTCTCCACCCAAGCGGCCGGAGAACCTGCGTGCAATCC
ACTGGGGGCGCG

Fig. 17

pCR-XL-TOPO-CMV-PUR-attB (SEQ ID NO: 10)

```
AGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGC
ACGACAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGC
TCACTCATTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAA
TTGTGAGCGGATAACAATTTCACACAGGAAACAGCTATGACCATGATTACGCCAAGCTAT
TTAGGTGACGCGTTAGAATACTCAAGCTATGCATCAAGCTTGGTACCGAGCTCGGATCCA
CTAGTAACGGCCGCCAGTGTGCTGGAATTCGCCCTTGGCCGCAATAAAATATCTTTATTT
TCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAGTACTAACATACGCTCTCCAT
CAAAACAAAACGAAACAAAACAAACTAGCAAAATAGGCTGTCCCCAGTGCAAGTGCAGGT
GCCAGAACATTTCTCTATCGATAGGTACCGAGCTCTTACGCGTGCTAGCCCTCGAGCAGG
ATCTATACATTGAATCAATATTGGCAATTAGCCATATTAGTCATTGGTTATATAGCATAA
ATCAATATTGGCTATTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTAT
ATTGGCTCATGTCCAATATGACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAG
TAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTT
ACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATG
ACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTAT
TTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCT
ATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGG
GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG
TTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC
CACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAA
TGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC
TATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGT
TTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCCCTCGAAGCTCGACTCTAGG
GGCTCGAGATCTGCGATCTAAGTAAGCTTGCATGCCTGCAGGTCGGCCGCCACGACCGGT
GCCGCCACCATCCCCTGACCCACGCCCCTGACCCCTCACAAGGAGACGACCTTCCATGAC
CGAGTACAAGCCCACGGTGCGCCTCGCCACCCGCGACGACGTCCCCCGGGCCGTACGCAC
CCTCGCCGCCGCGTTCGCCGACTACCCCGCCACGCGCCACACCGTCGACCCGGACCGCCA
CATCGAGCGGGTCACCGAGCTGCAAGAACTCTTCCTCACGCGCGTCGGGCTCGACATCGG
CAAGGTGTGGGTCGCGGACGACGGCGCCGCGGTGGCGGTCTGGACCACGCCGGAGAGCGT
CGAAGCGGGGGGCGGTGTTCGCCGAGATCGGCCCGCGCATGGCCGAGTTGAGCGGTTCCCG
GCTGGCCGCGCAGCAACAGATGGAAGGCCTCCTGGCGCCGCACCGGCCCAAGGAGCCCGC
GTGGTTCCTGGCCACCGTCGGCGTCTCGCCCGACCACCAGGGCAAGGGTCTGGGCAGCGC
CGTCGTGCTCCCCGGAGTGGAGGCGGCCGAGCGCGCCGGGGTGCCCGCCTTCCTGGAGAC
CTCCGCGCCCCGCAACCTCCCCTTCTACGAGCGGCTCGGCTTCACCGTCACCGCCGACGT
CGAGGTGCCCGAAGGACCGCGCACCTGGTGCATGACCCGCAAGCCCGGTGCCTGACGCCC
GCCCCACGACCCGCAGCGCCCGACCGAAAGGAGCGCACGACCCCATGGCTCCGACCGAAG
CCGACCCGGGCGGCCCCGCCGACCCCGCACCCGCCCCCGAGGCCCACCGACTCTAGAGTC
GGGGCGGCCGGCCGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCA
CAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTAT
TTGTAACCATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGT
TTCAGGTTCAGGGGGAGGTGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTG
GTAAAATCGATAAGGATCAATTCGGCTTCAGGTACCGTCGACGATGTAGGTCACGGTCTC
GAAGCCGCGGTGCGGGTGCCAGGGCGTGCCCTTGGGCTCCCCGGGCGCGTACTCCACCTC
ACCCATCTGGTCCATCATGATGAACGGGTCGAGGTGGCGGTAGTTGATCCCGGCGAACGC
GCGGCGCACCGGGAAGCCCTCGCCCTCGAAACCGCTGGGCGCGGTGGTCACGGTGAGCAC
GGGACGTGCGACGGCGTCGGCGGGTGCGGATACGCGGGGCAGCGTCAGCGGGTTCTCGAC
GGTCACGGCGGGCATGTCGACAGCCGAATTGATCCGTCGACCGATGCCCTTGAGAGCCTT
CAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGAC
TGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTTCCGCTTCCTCGC
TCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGG
CGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGAAGGGCGAAT
TCTGCAGATATCCATCACACTGGCGGCCGCTCGAGCATGCATCTAGAGGGCCCAATTCGC
CCTATAGTGAGTCGTATTACAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAA
ACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTA
```

ATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTATACGTACGGC
AGTTTAAGGTTTACACCTATAAAAGAGAGAGCCGTTATCGTCTGTTTGTGGATGTACAGA
GTGATATTATTGACACGCCGGGGCGACGGATGGTGATCCCCCTGGCCAGTGCACGTCTGC
TGTCAGATAAAGTCTCCCGTGAACTTTACCCGGTGGTGCATATCGGGGATGAAAGCTGGC
GCATGATGACCACCGATATGGCCAGTGTGCCGGTCTCCGTTATCGGGGAAGAAGTGGCTG
ATCTCAGCCACCGCGAAAATGACATCAAAAACGCCATTAACCTGATGTTCTGGGGAATAT
AAATGTCAGGCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTCACGTAGAAAG
CCAGTCCGCAGAAACGGTGCTGACCCCGGATGAATGTCAGCTACTGGGCTATCTGGACAA
GGGAAAACGCAAGCGCAAAGAGAAAGCAGGTAGCTTGCAGTGGGCTTACATGGCGATAGC
TAGACTGGGCGGTTTTATGGACAGCAAGCGAACCGGAATTGCCAGCTGGGGCGCCCTCTG
GTAAGGTTGGGAAGCCCTGCAAAGTAAACTGGATGGCTTTCTCGCCGCCAAGGATCTGAT
GGCGCAGGGGATCAAGCTCTGATCAAGAGACAGGATGAGGATCGTTTCGCATGATTGAAC
AAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCGGCTATGACT
GGGCACAACAGACAATCGGCTGCTCTGATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGC
GCCCGGTTCTTTTTGTCAAGACCGACCTGTCCGGTGCCCTGAATGAACTGCAAGACGAGG
CAGCGCGGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTG
TCACTGAAGCGGGAAGGGACTGGCTGCTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGT
CATCTCACCTTGCTCCTGCCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGC
ATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGAAACATCGCATCGAGCGAG
CACGTACTCGGATGGAAGCCGGTCTTGTCGATCAGGATGATCTGGACGAAGAGCATCAGG
GGCTCGCGCCAGCCGAACTGTTCGCCAGGCTCAAGGCGAGCATGCCCGACGGCGAGGATC
TCGTCGTGACCCATGGCGATGCCTGCTTGCCGAATATCATGGTGGAAAATGGCCGCTTTT
CTGGATTCATCGACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGG
CTACCCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTT
ACGGTATCGCCGCTCCCGATTCGCAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCT
TCTGAATTATTAACGCTTACAATTTCCTGATGCGGTATTTTCTCCTTACGCATCTGTGCG
GTATTTCACACCGCATACAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTG
TTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAAT
GCTTCAATAATAGCACGTGAGGAGGGCCACCATGGCCAAGTTGACCAGTGCCGTTCCGGT
GCTCACCGCGCGCGACGTCGCCGGAGCGGTCGAGTTCTGGACCGACCGGCTCGGGTTCTC
CCGGGACTTCGTGGAGGACGACTTCGCCGGTGTGGTCCGGGACGACGTGACCCTGTTCAT
CAGCGCGGTCCAGGACCAGGTGGTGCCGGACAACACCCTGGCCTGGGTGTGGGTGCGCGG
CCTGGACGAGCTGTACGCCGAGTGGTCGGAGGTCGTGTCCACGAACTTCCGGGACGCCTC
CGGGCCGGCCATGACCGAGATCGGCGAGCAGCCGTGGGGGCGGGAGTTCGCCCTGCGCGA
CCCGGCCGGCAACTGCGTGCACTTCGTGGCCGAGGAGCAGGACTGACACGTGCTAAAACT
TCATTTTTAATTTAAAAGGATCTAGGTGAAGATCCTTTTTGATAATCTCATGACCAAAAT
CCCTTAACGTGAGTTTTCGTTCCACTGAGCGTCAGACCCCGTAGAAAAGATCAAAGGATC
TTCTTGAGATCCTTTTTTTCTGCGCGTAATCTGCTGCTTGCAAACAAAAAAACCACCGCT
ACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCTACCAACTCTTTTTCCGAAGGTAACTGG
CTTCAGCAGAGCGCAGATACCAAATACTGTCCTTCTAGTGTAGCCGTAGTTAGGCCACCA
CTTCAAGAACTCTGTAGCACCGCCTACATACCTCGCTCTGCTAATCCTGTTACCAGTGGC
TGCTGCCAGTGGCGATAAGTCGTGTCTTACCGGGTTGGACTCAAGACGATAGTTACCGGA
TAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTCGTGCACACAGCCCAGCTTGGAGCGAAC
GACCTACACCGAACTGAGATACCTACAGCGTGAGCTATGAGAAAGCGCCACGCTTCCCGA
AGGGAGAAAGGCGGACAGGTATCCGGTAAGCGGCAGGGTCGGAACAGGAGAGCGCACGAG
GGAGCTTCCAGGGGGAAACGCCTGGTATCTTTATAGTCCTGTCGGGTTTCGCCACCTCTG
ACTTGAGCGTCGATTTTTGTGATGCTCGTCAGGGGGGCGGAGCCTATGGAAAAACGCCAG
CAACGCGGCCTTTTTACGGTTCCTGGGCTTTTGCTGGCCTTTTGCTCACATGTTCTTTCC
TGCGTTATCCCCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGC
TCGCCGCAGCCGAACGACCGAGCGCAGCGAGTCAGTGAGCGAGGAAGCGGAAG .

FIG. 18

attP (SEQ ID NO: 11)
GACTAGTACTGACGGACACACCGAAGCCCCGGCGGCAACCCTCAGCGGATGCCCCGGGGCTT
CACGTTTTCCCAGGTCAGAAGCGGTTTTCGGGAGTAGTGCCCCAACTGGGGTAACCTTTGAG
TTCTCTCAGTTGGGGGCGTAGGGTCGCCGACATGACACAAGGGGTTGTGACCGGGGTGGACA
CGTACGCGGGTGCTTACGACCGTCAGTCGCGCGAGCGCGACTAGTACA

Fig. 19
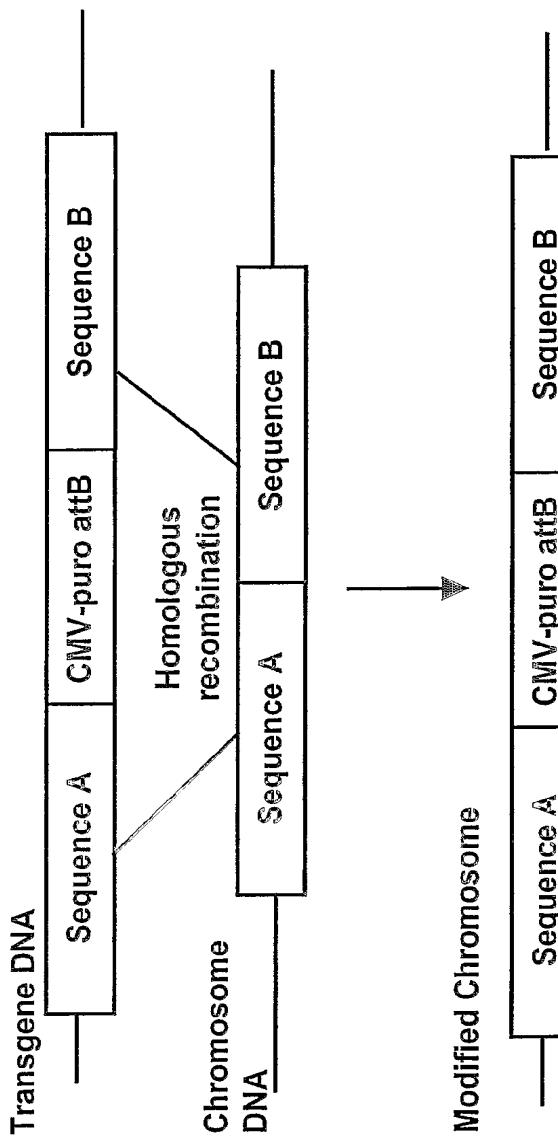
Fig. 20
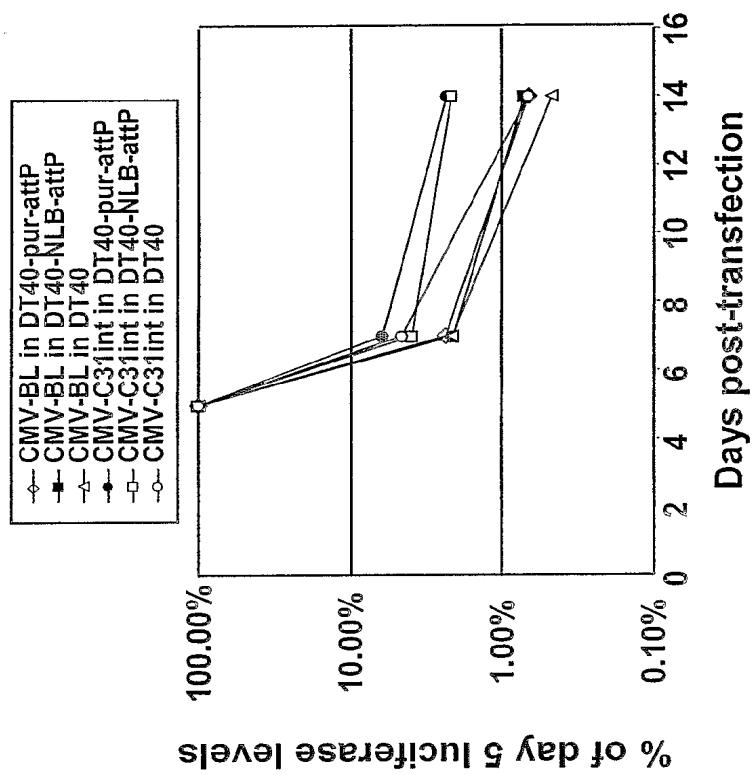
Fig. 21

Integrase.ST25
SEQUENCE LISTING

<110>   AviGenics, Inc

<120>   Avian Integrase-mediated Transformation

<130>   A181 1080.1

<160>   12

<170>   PatentIn version 3.2

<210>   1
<211>   6230
<212>   DNA
<213>   Plasmid pCMV-31int

<400>   1
cattcgccat tcaggctgcg caactgttgg gaagggcgat cggtgcgggc ctcttcgcta   60

ttacgccagc caatacgcaa accgcctctc cccgcgcgtt ggccgattca ttaatgcagg   120

atcgatccag acatgataag atacattgat gagtttggac aaaccacaac tagaatgcag   180

tgaaaaaaat gctttatttg tgaaatttgt gatgctattg ctttatttgt aaccattata   240

agctgcaata aacaagttaa caacaacaat tgcattcatt ttatgtttca ggttcagggg   300

gaggtgtggg aggttttttta aagcaagtaa aacctctaca aatgtggtat ggctgattat   360

gatcatgaac agactgtgag gactgagggg cctgaaatga gccttgggac tgtgaatcta   420

aaatacacaa acaattagaa tcactagctc ctgtgtataa tattttcata aatcatactc   480

agtaagcaaa actctcaagc agcaagcata tgcagctagt ttaacacatt atacacttaa   540

aaattttata tttaccttag agctttaaat ctctgtaggt agtttgtcca attatgtcac   600

accacagaag taaggttcct tcacaaagat cccaagctag cttataatac gactcactat   660

agggagagag ctatgacgtc gcatgcacgc gtaagcttgg gcccctcgag ggatccgggt   720

gtctcgctac gccgctacgt cttccgtgcc gtcctgggcg tcgtcttcgt cgtcgtcggt   780

cggcggcttc gcccacgtga tcgaagcgcg cttctcgatg ggcgttccct gccccctgcc   840

cgtagtcgac ttcgtgacaa cgatcttgtc tacgaagagc ccgacgaaca cgcgcttgtc   900

gtctactgac gcgcgccccc accacgactt agggccggtc gggtcagcgt cggcgtcttc   960

ggggaaccat tggtcaaggg gaagcttcgg ggcttcggcg gcttcaagtt cggcaagccg   1020

ctcttccgcc ccttgctgcc ggagcgtcag cgctgcctgt tgcttccgga agtgcttcct   1080

gccaacgggt ccgtcgtacg cgcctgccgc gcggtcttcg tacagctctt caagggcgtt   1140

cagggcgtcg gcgcgctccg caacaaggtt cgccgttcg ccgctcttct caggcgcctc   1200

agtgagcttg ccgaagcgtc gggcggcttc ccacagaagc gccaacgtct cttcgtcgcc   1260

ttcggcgtgc ctgatcttgt tgaagatgcg ttccgcaacg aacttgtcga gtgccgccat   1320

gctgacgttg cacgtgcctt cgtgctgccc aggtgcggac gggtcgacca ccttccggcg   1380

acggcagcgg taagagtcct tgatcgattc ttccccgcgc ttcgaagtca tgacggcgcc   1440

acactcgcag tacagcttgt ccatggcgga cagaatggct tgcccccggg aaagcccctt   1500

Page 1

Integrase.ST25

```
gccgcgcccc ctgccgtcca accacgcctg aagctcatac cactcagcgg gctcgatgat   1560
cggtccgcaa tcaagctcga ccggccggag cgtgatcggg tcgcgctgaa tgcggtaacc   1620
ctcaatcttc gtggtcggcg tgccgtccgg cttcttcttg tagatcacct cagcggcgaa   1680
gcccgcaata cgcgggtccc gaaggattcg cataacggtt gccgggtccc aggcgcttga   1740
agcggtcttc ttcccaatcg tctcgccccg ggtcggcacg gcgtcagcgt ccatgcgctt   1800
acaaagcccc gtgatgctgc ccgggtgaat ggcggcttga ctgcccggct tgaagggaag   1860
gtgtttgtgc gtcttgatct cacgccacca ccaccggatt acgtcgggct cgaactcgaa   1920
gggtccggta aggggagtgg tcgagtgcgc aagcttgttg atgacgacat tgaccattcg   1980
gccgttgcgc gtgatctcct tcgtctccga aacaagctcg aagccgtaag gcgccttccc   2040
gccgacgtac ccgcccaatt cgcgctgaag gttcttcgtg tcgagaatct tcgccgactt   2100
cagcgaagat tctttgtgcg acgcgtcgag ccgcataatc aggtgaatca ggtccatgac   2160
gtttccctgc cggaagacgc cttcctgagt ggaaacaatc gtcacgccca gggcgagcaa   2220
ttccgagaca atcggaatcg cgtccatgac cttcaggcgc gagaagcgcg acacgtcata   2280
gacaatgatc atgttgagcc gcccggcgcg gcattcgttc aggatgcgtt cgaactccgg   2340
gcgctccgcc gtcccgaacg ccgacgtgcc cggcgcttcg ctgaaatgcc cgacgaacct   2400
gaaccggccc ccgtcgcgct cgacttcgcg ctgaaggtcg gccgccttgt cttcgttggc   2460
gctacgctgt gtcgctgggc ttgctgcgct cgaattctcg cgctcgcgcg actgacggtc   2520
gtaagcaccc gcgtacgtgt ccacccggt cacaacccct tgtgtcatgt cggcgaccct   2580
acgactagtg agctcgtcga cccgggaatt ccggaccggt acctgcaggc gtaccttcta   2640
tagtgtcacc taaatagctt tttgcaaaag cctaggctag agtccggagg ctggatcggt   2700
cccggtgtct tctatggagg tcaaaacagc gtggatggcg tctccaggcg atctgacggt   2760
tcactaaacg agctctgctt atatagacct cccaccgtac acgcctaccg cccatttgcg   2820
tcaatggggc ggagttgtta cgacattttg gaaagtcccg ttgattttgg tgccaaaaca   2880
aactcccatt gacgtcaatg gggtggagac ttggaaatcc ccgtgagtca aaccgctatc   2940
cacgcccatt gatgtactgc caaaaccgca tcaccatggt aatagcgatg actaatacgt   3000
agatgtactg ccaagtagga aagtcccata aggtcatgta ctgggcataa tgccaggcgg   3060
gccatttacc gtcattgacg tcaataggg gcgtacttgg catatgatac acttgatgta   3120
ctgccaagtg ggcagtttac cgtaaatact ccaccattg acgtcaatgg aaagtcccta   3180
ttggcgttac tatgggaaca tacgtcatta ttgacgtcaa tgggcggggg tcgttgggcg   3240
gtcagccagg cgggccattt accgtaagtt atgtaacgac ctgcacgatg ctgtttcctg   3300
tgtgaaattg ttatccgctc acaattccac acattatacg agccggaagc tataaagtgt   3360
aaagcctggg gtgcctaatg agtgaaaggg cctcgtatac gcctattttt ataggttaat   3420
gtcatgataa taatggtttc ttagacgtca ggtggcactt ttcggggaaa tgtgcgcgga   3480
accctatttt gtttattttt ctaaatacat tcaaatatgt atccgctcat gagacaataa   3540
```

Integrase.ST25

```
ccctgataaa tgcttcaata atattgaaaa acgcgcgaat tgcaagctct gcattaatga      3600
atcggccaac gcgcgggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc      3660
actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg      3720
gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc      3780
cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctccgc      3840
cccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga      3900
ctataaagat accaggcgtt cccccctgga agctccctcg tgcgctctcc tgttccgacc      3960
ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcaa      4020
tgctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg      4080
cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc      4140
aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga      4200
gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact      4260
agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt      4320
ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggttttt tgtttgcaag      4380
cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg      4440
tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgcc ataacttcgt      4500
atagcataca ttatacgaag ttatggcatg agattatcaa aaaggatctt cacctagatc      4560
ctttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct      4620
gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca      4680
tccatagttg cctgactccc cgtcgtgtag ataactacga tacgggaggg cttaccatct      4740
ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca      4800
ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc      4860
atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg      4920
cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct      4980
tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgcaaa      5040
aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta      5100
tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc      5160
ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg      5220
agttgctctt gcccggcgtc aatacgggat aataccgcgc cacatagcag aactttaaaa      5280
gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg      5340
agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc      5400
accagcgttt ctgggtgagc aaaaacagga aggcaaaatg ccgcaaaaaa gggaataagg      5460
gcgacacgga aatgttgaat actcatactc ttcctttttc aatattattg aagcatttat      5520
cagggttatt gtctcatgcc aggggtgggc acacatattt gataccagcg atccctacac      5580
```

```
                              Integrase.ST25
agcacataat tcaatgcgac ttccctctat cgcacatctt agacctttat tctccctcca    5640

gcacacatcg aagctgccga gcaagccgtt ctcaccagtc caagacctgg catgagcgga    5700

tacatatttg aatgtattta gaaaaataaa caaataggg ttccgcgcac atttccccga    5760

aaagtgccac ctgaaattgt aaacgttaat attttgttaa aattcgcgtt aaattttttgt   5820

taaatcagct catttttttaa ccaataggcc gaaatcggca aaatccctta taaatcaaaa    5880

gaatagaccg agatagggtt gagtgttgtt ccagtttgga acaagagtcc actattaaag    5940

aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg cccactacgt    6000

gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc gtaaagcact aaatcggaac    6060

cctaaaggga gccccgatt tagagcttga cggggaaagc cggcgaacgt ggcgagaaag    6120

gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc ggtcacgctg    6180

cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac agggcgcgtc              6230


<210>    2
<211>    5982
<212>    DNA
<213>    Plasmid pCMV-luc-attB

<400>    2
ctctatcgat aggtaccgag ctcttacgcg tgctagccct cgagcaggat ctatacattg      60

aatcaatatt ggcaattagc catattagtc attggttata tagcataaat caatattggc     120

tattggccat tgcatacgtt gtatctatat cataatatgt acatttatat tggctcatgt     180

ccaatatgac cgccatgttg acattgatta ttgactagtt attaatagta atcaattacg     240

gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc     300

ccgcctggct gaccgcccaa cgaccccgc ccattgacgt caataatgac gtatgttccc     360

atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt acggtaaact     420

gcccacttgg cagtacatca agtgtatcat atgccaagtc cgccccctat tgacgtcaat     480

gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttacggga ctttcctact     540

tggcagtaca tctacgtatt agtcatcgct attaccatgg tgatgcggtt ttggcagtac     600

atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca ccccattgac     660

gtcaatggga gtttgttttg gcaccaaaat caacgggact ttccaaaatg tcgtaacaac     720

tccgccccat tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga     780

gctcgtttag tgaaccgtca gatcgcctgg agacgccatc cacgctgttt tgacctccat     840

agaagacacc gggaccgatc cagcctcccc tcgaagctcg actctagggg ctcgagatct     900

gcgatctaag taagcttggc attccggtac tgttggtaaa gccaccatgg aagacgccaa     960

aaacataaag aaaggcccgg cgccattcta tccgctggaa gatggaaccg ctggagagca    1020

actgcataag gctatgaaga gatacgccct ggttcctgga caattgcttt ttacagatgc    1080

acatatcgag gtggacatca cttacgctga gtacttcgaa atgtccgttc ggttggcaga    1140

agctatgaaa cgatatgggc tgaatacaaa tcacagaatc gtcgtatgca gtgaaaactc    1200
```

Page 4

Integrase.ST25

```
tcttcaattc tttatgccgg tgttgggcgc gttatttatc ggagttgcag ttgcgcccgc    1260
gaacgacatt tataatgaac gtgaattgct caacagtatg ggcatttcgc agcctaccgt    1320
ggtgttcgtt tccaaaaagg ggttgcaaaa aattttgaac gtgcaaaaaa agctcccaat    1380
catccaaaaa attattatca tggattctaa aacggattac cagggatttc agtcgatgta    1440
cacgttcgtc acatctcatc tacctcccgg ttttaatgaa tacgattttg tgccagagtc    1500
cttcgatagg gacaagacaa ttgcactgat catgaactcc tctggatcta ctggtctgcc    1560
taaaggtgtc gctctgcctc atagaactgc ctgcgtgaga ttctcgcatg ccagagatcc    1620
tattttttggc aatcaaatca ttccggatac tgcgatttta agtgttgttc cattccatca    1680
cggttttgga atgtttacta cactcggata tttgatatgt ggatttcgag tcgtcttaat    1740
gtatagattt gaagaagagc tgtttctgag gagccttcag gattacaaga ttcaaagtgc    1800
gctgctggtg ccaaccctat tctccttctt cgccaaaagc actctgattg acaaatacga    1860
tttatctaat ttacacgaaa ttgcttctgg tggcgctccc ctctctaagg aagtcgggga    1920
agcggttgcc aagaggttcc atctgccagg tatcaggcaa ggatatgggc tcactgagac    1980
tacatcagct attctgatta cacccgaggg ggatgataaa ccgggcgcgg tcggtaaagt    2040
tgttccattt tttgaagcga aggttgtgga tctggatacc gggaaaacgc tgggcgttaa    2100
tcaaagaggc gaactgtgtg tgagaggtcc tatgattatg tccggttatg taaacaatcc    2160
ggaagcgacc aacgccttga ttgacaagga tggatggcta cattctggag acatagctta    2220
ctgggacgaa gacgaacact tcttcatcgt tgaccgcctg aagtctctga ttaagtacaa    2280
aggctatcag gtggctcccg ctgaattgga atccatcttg ctccaacacc ccaacatctt    2340
cgacgcaggt gtcgcaggtc ttcccgacga tgacgccggt gaacttcccg ccgccgttgt    2400
tgtttttggag cacggaaaga cgatgacgga aaaagagatc gtggattacg tcgccagtca    2460
agtaacaacc gcgaaaaagt tgcgcggagg agttgtgttt gtggacgaag taccgaaagg    2520
tcttaccgga aaactcgacg caagaaaaat cagagagatc ctcataaagg ccaagaaggg    2580
cggaaagatc gccgtgtaat tctagagtcg gggcggccgg ccgcttcgag cagacatgat    2640
aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa aatgctttat    2700
ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca ataaacaagt    2760
taacaacaac aattgcattc attttatgtt tcaggttcag ggggaggtgt gggaggtttt    2820
ttaaagcaag taaaacctct acaaatgtgg taaaatcgat aaggatcaat tcggcttcag    2880
gtaccgtcga cgatgtaggt cacggtctcg aagccgcggt gcgggtgcca gggcgtgccc    2940
ttgggctccc cgggcgcgta ctccacctca cccatctggt ccatcatgat gaacgggtcg    3000
aggtggcggt agttgatccc ggcgaacgcg cggcgcaccg gaagccctc gccctcgaaa     3060
ccgctgggcg cggtggtcac ggtgagcacg gacgtgcga cggcgtcggc gggtgcggat     3120
acgcggggca gcgtcagcgg ttctcgacg tcacggcgg gcatgtcgac agccgaattg      3180
atccgtcgac cgatgccctt gagagccttc aacccagtca gctccttccg gtgggcgcgg    3240
```

Integrase.ST25

```
ggcatgacta tcgtcgccgc acttatgact gtcttcttta tcatgcaact cgtaggacag   3300

gtgccggcag cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct   3360

gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga   3420

taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc   3480

cgcgttgctg gcgtttttcc ataggctccg ccccccctgac gagcatcaca aaaatcgacg   3540

ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg   3600

aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt   3660

tctcccttcg ggaagcgtgg cgctttctca atgctcacgc tgtaggtatc tcagttcggt   3720

gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg   3780

cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact   3840

ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt   3900

cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct   3960

gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca acaaaccac    4020

cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc   4080

tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg   4140

ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta   4200

aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca   4260

atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc   4320

ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc   4380

tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc   4440

agccggaagg gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat   4500

taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt   4560

tgccattgct acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc   4620

cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa aagcggttag   4680

ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt   4740

tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac   4800

tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg   4860

cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat   4920

tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga tccagttc    4980

gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc   5040

tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag gaataaggg cgacacggaa    5100

atgttgaata ctcatactct tcctttttca atattattga agcatttatc agggttattg   5160

tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg   5220

cacatttccc cgaaaagtgc cacctgacgc gccctgtagc ggcgcattaa gcgcggcggg   5280
```

Page 6

Integrase.ST25

```
tgtggtggtt acgcgcagcg tgaccgctac acttgccagc gccctagcgc ccgctccttt    5340

cgctttcttc ccttcctttc tcgccacgtt cgccggcttt ccccgtcaag ctctaaatcg    5400

ggggctccct ttagggttcc gatttagtgc tttacggcac ctcgacccca aaaaacttga    5460

ttagggtgat ggttcacgta gtgggccatc gccctgatag acggtttttc gccctttgac    5520

gttggagtcc acgttcttta atagtggact cttgttccaa actggaacaa cactcaaccc    5580

tatctcggtc tattcttttg atttataagg gattttgccg atttcggcct attggttaaa    5640

aaatgagctg atttaacaaa aatttaacgc gaattttaac aaaatattaa cgtttacaat    5700

ttcccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc    5760

gctattacgc cagcccaagc taccatgata agtaagtaat attaaggtac gggaggtact    5820

tggagcggcc gcaataaaat atctttattt tcattacatc tgtgtgttgg ttttttgtgt    5880

gaatcgatag tactaacata cgctctccat caaaacaaaa cgaaacaaaa caaactagca    5940

aaataggctg tccccagtgc aagtgcaggt gccagaacat tt                       5982
```

<210> 3
<211> 5924
<212> DNA
<213> Plasmid pCMV-luc-attP

<400> 3

```
ctctatcgat aggtaccgag ctcttacgcg tgctagccct cgagcaggat ctatacattg    60

aatcaatatt ggcaattagc catattagtc attggttata tagcataaat caatattggc    120

tattggccat tgcatacgtt gtatctatat cataatatgt acatttatat tggctcatgt    180

ccaatatgac cgccatgttg acattgatta ttgactagtt attaatagta atcaattacg    240

gggtcattag ttcatagccc atatatggag ttccgcgtta cataacttac ggtaaatggc    300

ccgcctggct gaccgcccaa cgacccccgc ccattgacgt caataatgac gtatgttccc    360

atagtaacgc caatagggac tttccattga cgtcaatggg tggagtattt acggtaaact    420

gcccacttgg cagtacatca agtgtatcat atgccaagtc cgcccctat tgacgtcaat    480

gacggtaaat ggcccgcctg gcattatgcc cagtacatga ccttacggga ctttcctact    540

tggcagtaca tctacgtatt agtcatcgct attaccatgg tgatgcggtt ttggcagtac    600

atcaatgggc gtggatagcg gtttgactca cggggatttc caagtctcca ccccattgac    660

gtcaatggga gtttgttttg gcaccaaaat caacgggact ttccaaaatg tcgtaacaac    720

tccgccccat tgacgcaaat gggcggtagg cgtgtacggt gggaggtcta tataagcaga    780

gctcgtttag tgaaccgtca gatcgcctgg agacgccatc cacgctgttt tgacctccat    840

agaagacacc gggaccgatc cagcctcccc tcgaagctcg actctagggg ctcgagatct    900

gcgatctaag taagcttggc attccggtac tgttggtaaa gccaccatgg aagacgccaa    960

aaacataaag aaaggcccgg cgccattcta ccgctggaa gatggaaccg ctggagagca    1020

actgcataag gctatgaaga gatacgccct ggttcctgga acaattgctt ttacagatgc    1080
```

Integrase.ST25

```
acatatcgag gtggacatca cttacgctga gtacttcgaa atgtccgttc ggttggcaga    1140

agctatgaaa cgatatgggc tgaatacaaa tcacagaatc gtcgtatgca gtgaaaactc    1200

tcttcaattc tttatgccgg tgttgggcgc gttatttatc ggagttgcag ttgcgcccgc    1260

gaacgacatt tataatgaac gtgaattgct caacagtatg ggcatttcgc agcctaccgt    1320

ggtgttcgtt ccaaaaagg ggttgcaaaa aattttgaac gtgcaaaaaa agctcccaat    1380

catccaaaaa attattatca tggattctaa aacggattac cagggatttc agtcgatgta    1440

cacgttcgtc acatctcatc tacctcccgg ttttaatgaa tacgattttg tgccagagtc    1500

cttcgatagg gacaagacaa ttgcactgat catgaactcc tctggatcta ctggtctgcc    1560

taaaggtgtc gctctgcctc atagaactgc ctgcgtgaga ttctcgcatg ccagagatcc    1620

tattttttggc aatcaaatca ttccggatac tgcgatttta agtgttgttc cattccatca    1680

cggttttgga atgtttacta cactcggata tttgatatgt ggatttcgag tcgtcttaat    1740

gtatagattt gaagaagagc tgtttctgag gagccttcag gattacaaga ttcaaagtgc    1800

gctgctggtg ccaaccctat tctccttctt cgccaaaagc actctgattg acaaatacga    1860

tttatctaat ttacacgaaa ttgcttctgg tggcgctccc ctctctaagg aagtcgggga    1920

agcggttgcc aagaggttcc atctgccagg tatcaggcaa ggatatgggc tcactgagac    1980

tacatcagct attctgatta cacccgaggg ggatgataaa ccgggcgcgg tcggtaaagt    2040

tgttccattt tttgaagcga aggttgtgga tctggatacc gggaaaacgc tgggcgttaa    2100

tcaaagaggc gaactgtgtg tgagaggtcc tatgattatg tccggttatg taaacaatcc    2160

ggaagcgacc aacgccttga ttgacaagga tggatggcta cattctggag acatagctta    2220

ctgggacgaa gacgaacact tcttcatcgt tgaccgcctg aagtctctga ttaagtacaa    2280

aggctatcag gtggctcccg ctgaattgga atccatcttg ctccaacacc ccaacatctt    2340

cgacgcaggt gtcgcaggtc ttcccgacga tgacgccggt gaacttcccg ccgccgttgt    2400

tgttttggag cacggaaaga cgatgacgga aaaagagatc gtggattacg tcgccagtca    2460

agtaacaacc gcgaaaaagt tgcgcggagg agttgtgttt gtggacgaag taccgaaagg    2520

tcttaccgga aaactcgacg caagaaaaat cagagagatc ctcataaagg ccaagaaggg    2580

cggaaagatc gccgtgtaat ctagagtcg gggcggccgg ccgcttcgag cagacatgat    2640

aagatacatt gatgagtttg gacaaaccac aactagaatg cagtgaaaaa aatgctttat    2700

ttgtgaaatt tgtgatgcta ttgctttatt tgtaaccatt ataagctgca ataaacaagt    2760

taacaacaac aattgcattc attttatgtt tcaggttcag ggggaggtgt gggaggtttt    2820

ttaaagcaag taaaacctct acaaatgtgg taaaatcgat aaggatcaat tcggcttcga    2880

ctagtactga cggacacacc gaagccccgg cggcaaccct cagcggatgc cccggggctt    2940

cacgttttcc caggtcagaa gcggttttcg ggagtagtgc cccaactggg gtaacctttg    3000

agttctctca gttggggggcg tagggtcgcc gacatgacac aaggggttgt gaccggggtg    3060

gacacgtacg cgggtgctta cgaccgtcag tcgcgcgagc gcgactagta caagccgaat    3120
```

Page 8

Integrase.ST25

```
tgatccgtcg accgatgccc ttgagagcct tcaacccagt cagctccttc cggtgggcgc    3180

ggggcatgac tatcgtcgcc gcacttatga ctgtcttctt tatcatgcaa ctcgtaggac    3240

aggtgccggc agcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg    3300

ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg    3360

gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag    3420

gccgcgttgc tggcgttttt ccataggctc cgcccccctg acgagcatca caaaaatcga    3480

cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct    3540

ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc    3600

tttctccctt cgggaagcgt ggcgctttct caatgctcac gctgtaggta tctcagttcg    3660

gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc    3720

tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca    3780

ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag    3840

ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct    3900

ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc    3960

accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga    4020

tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa cgaaaactca    4080

cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaaat    4140

taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac    4200

caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt    4260

gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt    4320

gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag    4380

ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct    4440

attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt    4500

gttgccattg ctacaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc    4560

tccggttccc aacgatcaag cgagttaca tgatccccca tgttgtgcaa aaaagcggtt    4620

agctccttcg gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg    4680

gttatggcag cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg    4740

actggtgagt actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct    4800

tgcccggcgt caatacggga taataccgcg ccacatagca gaactttaaa agtgctcatc    4860

attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt gagatccagt    4920

tcgatgtaac ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt    4980

tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa agggaataag gcgacacgg    5040

aaatgttgaa tactcatact cttccttttt caatattatt gaagcattta tcagggttat    5100

tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat aggggttccg    5160
```

Page 9

```
                              Integrase.ST25
cgcacatttc cccgaaaagt gccacctgac gcgccctgta gcggcgcatt aagcgcggcg     5220

ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgccctagc gcccgctcct     5280

ttcgcttttct tcccttcctt tctcgccacg ttcgccggct ttccccgtca agctctaaat     5340

cggggggctcc ctttagggtt ccgatttagt gctttacggc acctcgaccc caaaaaactt     5400

gattagggtg atggttcacg tagtgggcca tcgccctgat agacggtttt tcgccctttg     5460

acgttggagt ccacgttctt taatagtgga ctcttgttcc aaactggaac aacactcaac     5520

cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta     5580

aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgtttaca     5640

atttcccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct     5700

tcgctattac gccagcccaa gctaccatga taagtaagta atattaaggt acgggaggta     5760

cttggagcgg ccgcaataaa atatctttat tttcattaca tctgtgtgtt ggttttttgt     5820

gtgaatcgat agtactaaca tacgctctcc atcaaaacaa aacgaaacaa aacaaactag     5880

caaaataggc tgtccccagt gcaagtgcag gtgccagaac attt                      5924


<210>  4
<211>  5101
<212>  DNA
<213>  Plasmid pCMV-pur-attB

<400>  4
ctagagtcgg ggcggccggc cgcttcgagc agacatgata agatacattg atgagtttgg      60

acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat     120

tgctttattt gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca     180

ttttatgttt caggttcagg gggaggtgtg ggaggttttt taaagcaagt aaaacctcta     240

caaatgtggt aaaatcgata aggatcaatt cggcttcagg taccgtcgac gatgtaggtc     300

acggtctcga agccgcggtg cgggtgccag ggcgtgccct tgggctcccc gggcgcgtac     360

tccacctcac ccatctggtc catcatgatg aacgggtcga ggtggcggta gttgatcccg     420

gcgaacgcgc ggcgcaccgg gaagccctcg ccctcgaaac cgctgggcgc ggtggtcacg     480

gtgagcacgg gacgtgcgac ggcgtcggcg ggtgcggata cgcggggcag cgtcagcggg     540

ttctcgacgg tcacggcggg catgtcgaca gccgaattga tccgtcgacc gatgcccttg     600

agagccttca acccagtcag ctccttccgg tgggcgcggg gcatgactat cgtcgccgca     660

cttatgactg tcttctttat catgcaactc gtaggacagg tgccggcagc gctcttccgc     720

ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca     780

ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg     840

agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgttttttcca     900

taggctccgc cccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa     960

cccgacagga ctataaagat accaggcgtt ccccctggga agctccctcg tgcgctctcc    1020

tgttccgacc ctgccgcttta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc    1080
```

Integrase.ST25

```
gctttctcaa tgctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct   1140

gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg   1200

tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag   1260

gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta   1320

cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg   1380

aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt   1440

tgtttgcaag cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt   1500

ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag   1560

attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat   1620

ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc   1680

tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat   1740

aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc   1800

acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag   1860

aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag   1920

agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt   1980

ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg   2040

agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt   2100

tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc   2160

tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc   2220

attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa   2280

taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg   2340

aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc   2400

caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag   2460

gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt   2520

cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt   2580

tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc   2640

acctgacgcg ccctgtagcg gcgcattaag cgcggcgggt gtggtggtta cgcgcagcgt   2700

gaccgctaca cttgccagcg ccctagcgcc cgctcctttc gctttcttcc cttcctttct   2760

cgccacgttc gccggctttc cccgtcaagc tctaaatcgg gggctccctt tagggttccg   2820

atttagtgct ttacggcacc tcgaccccaa aaaacttgat tagggtgatg gttcacgtag   2880

tgggccatcg ccctgataga cggttttttcg ccctttgacg ttggagtcca cgttctttaa   2940

tagtggactc ttgttccaaa ctggaacaac actcaaccct atctcggtct attctttttga   3000

tttataaggg attttgccga tttcggccta ttggttaaaa aatgagctga tttaacaaaa   3060

atttaacgcg aattttaaca aaatattaac gtttacaatt tcccattcgc cattcaggct   3120
```

Page 11

Integrase.ST25

```
gcgcaactgt tgggaagggc gatcggtgcg ggcctcttcg ctattacgcc agcccaagct    3180
accatgataa gtaagtaata ttaaggtacg ggaggtactt ggagcggccg caataaaata    3240
tctttatttt cattacatct gtgtgttggt tttttgtgtg aatcgatagt actaacatac    3300
gctctccatc aaaacaaaac gaaacaaaac aaactagcaa aataggctgt ccccagtgca    3360
agtgcaggtg ccagaacatt tctctatcga taggtaccga gctcttacgc gtgctagccc    3420
tcgagcagga tctatacatt gaatcaatat tggcaattag ccatattagt cattggttat    3480
atagcataaa tcaatattgg ctattggcca ttgcatacgt tgtatctata tcataatatg    3540
tacatttata ttggctcatg tccaatatga ccgccatgtt gacattgatt attgactagt    3600
tattaatagt aatcaattac ggggtcatta gttcatagcc catatatgga gttccgcgtt    3660
acataactta cggtaaatgg cccgcctggc tgaccgccca acgacccccg cccattgacg    3720
tcaataatga cgtatgttcc catagtaacg ccaataggga ctttccattg acgtcaatgg    3780
gtggagtatt tacggtaaac tgcccacttg gcagtacatc aagtgtatca tatgccaagt    3840
ccgcccccta ttgacgtcaa tgacggtaaa tggcccgcct ggcattatgc ccagtacatg    3900
accttacggg actttcctac ttggcagtac atctacgtat tagtcatcgc tattaccatg    3960
gtgatgcggt tttggcagta catcaatggg cgtggatagc ggtttgactc acggggattt    4020
ccaagtctcc accccattga cgtcaatggg agtttgtttt ggcaccaaaa tcaacgggac    4080
tttccaaaat gtcgtaacaa ctccgcccca ttgacgcaaa tgggcggtag cgtgtacgg    4140
tgggaggtct atataagcag agctcgttta gtgaaccgtc agatcgcctg gagacgccat    4200
ccacgctgtt ttgacctcca tagaagacac cgggaccgat ccagcctccc ctcgaagctc    4260
gactctaggg gctcgagatc tgcgatctaa gtaagcttgc atgcctgcag gtcggccgcc    4320
acgaccggtg ccgccaccat cccctgaccc acgcccctga cccctcacaa ggagacgacc    4380
ttccatgacc gagtacaagc ccacggtgcg cctcgccacc cgcgacgacg tccccggggc    4440
cgtacgcacc ctcgccgccg cgttcgccga ctaccccgcc acgcgccaca ccgtcgaccc    4500
ggaccgccac atcgagcggg tcaccgagct gcaagaactc ttcctcacgc gcgtcgggct    4560
cgacatcggc aaggtgtggg tcgcggacga cggcgccgcg gtggcggtct ggaccacgcc    4620
ggagagcgtc gaagcggggg cggtgttcgc cgagatcggc ccgcgcatgg ccgagttgag    4680
cggttcccgg ctggccgcgc agcaacagat ggaaggcctc ctggcgccgc accggcccaa    4740
ggagcccgcg tggttcctgg ccaccgtcgg cgtctcgccc gaccaccagg gcaagggtct    4800
gggcagcgcc gtcgtgctcc cggagtgga ggcggccgag cgcgccgggg tgcccgcctt    4860
cctggagacc tccgcgcccc gcaacctccc cttctacgag cggctcggct tcaccgtcac    4920
cgccgacgtc gaggtgcccg aaggaccgcg cacctggtgc atgacccgca gcccggtgc    4980
ctgacgcccg ccccacgacc cgcagcgccc gaccgaaagg agcgcacgac cccatggctc    5040
cgaccgaagc cgacccgggc ggccccgccg accccgcacc cgcccccgag cccaccgac    5100
t                                                                      5101
```

Page 12

Integrase.ST25

<210>   5
<211>   5043
<212>   DNA
<213>   Plasmid pCMV-pur-attP

<400>   5
ctagagtcgg ggcggccggc cgcttcgagc agacatgata agatacattg atgagtttgg   60

acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat   120

tgctttattt gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca   180

ttttatgttt caggttcagg gggaggtgtg ggaggttttt taaagcaagt aaaacctcta   240

caaatgtggt aaaatcgata aggatcaatt cggcttcgac tagtactgac ggacacaccg   300

aagccccggc ggcaaccctc agcggatgcc ccggggcttc acgttttccc aggtcagaag   360

cggttttcgg gagtagtgcc ccaactgggg taacctttga gttctctcag ttggggggcgt   420

agggtcgccg acatgacaca aggggttgtg accggggtgg acacgtacgc gggtgcttac   480

gaccgtcagt cgcgcgagcg cgactagtac aagccgaatt gatccgtcga ccgatgccct   540

tgagagcctt caacccagtc agctccttcc ggtgggcgcg gggcatgact atcgtcgccg   600

cacttatgac tgtcttcttt atcatgcaac tcgtaggaca ggtgccggca gcgctcttcc   660

gcttcctcgc tcactgactc gctgcgctcg tcgttcggc tgcggcgagc ggtatcagct   720

cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg   780

tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgttttttc   840

cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga   900

aacccgacag gactataaag ataccaggcg tttcccctg gaagctccct cgtgcgctct   960

cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg   1020

gcgctttctc aatgctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag   1080

ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat   1140

cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac   1200

aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac   1260

tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc   1320

ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt   1380

tttgtttgca agcagcagat tacgcgcaga aaaaaggat ctcaagaaga tcctttgatc   1440

ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg   1500

agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca   1560

atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca   1620

cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag   1680

ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac   1740

ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc   1800

Page 13

Integrase.ST25

```
agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct   1860

agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc   1920

gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg   1980

cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc   2040

gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat   2100

tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag   2160

tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat   2220

aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg   2280

cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca   2340

cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga   2400

aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc   2460

ttcctttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata   2520

tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc cgaaaagtg   2580

ccacctgacg cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt tacgcgcagc   2640

gtgaccgcta cacttccag cgccctagcg cccgctcctt tcgctttctt cccttccttt   2700

ctcgccacgt tcgccggctt tccccgtcaa gctctaaatc gggggctccc tttagggttc   2760

cgatttagtg ctttacggca cctcgacccc aaaaaacttg attagggtga tggttcacgt   2820

agtgggccat cgccctgata gacggttttt cgccctttga cgttggagtc cacgttcttt   2880

aatagtggac tcttgttcca aactggaaca cactcaacc ctatctcggt ctattctttt   2940

gatttataag ggattttgcc gatttcggcc tattggttaa aaaatgagct gatttaacaa   3000

aaatttaacg cgaattttaa caaaatatta cgtttacaa tttcccattc gccattcagg   3060

ctgcgcaact gttgggaagg cgatcggtg cgggcctctt cgctattacg ccagcccaag   3120

ctaccatgat aagtaagtaa tattaaggta cgggaggtac ttggagcggc cgcaataaaa   3180

tatctttatt ttcattacat ctgtgtgttg gttttttgtg tgaatcgata gtactaacat   3240

acgctctcca tcaaaacaaa acgaaacaaa acaaactagc aaaataggct gtccccagtg   3300

caagtgcagg tgccagaaca tttctctatc gataggtacc gagctcttac gcgtgctagc   3360

cctcgagcag gatctataca ttgaatcaat attggcaatt agccatatta gtcattggtt   3420

atatagcata aatcaatatt ggctattggc cattgcatac gttgtatcta tatcataata   3480

tgtacattta tattggctca tgtccaatat gaccgccatg ttgacattga ttattgacta   3540

gttattaata gtaatcaatt acggggtcat tagttcatag cccatatatg gagttccgcg   3600

ttacataact tacggtaaat ggcccgcctg gctgaccgcc aacgacccc gcccattga   3660

cgtcaataat gacgtatgtt cccatagtaa cgccaatagg gactttccat tgacgtcaat   3720

gggtggagta tttacggtaa actgcccact ggcagtaca tcaagtgtat catatgccaa   3780

gtccgccccc tattgacgtc aatgacggta aatggcccgc ctggcattat gcccagtaca   3840
```

Page 14

Integrase.ST25

```
tgaccttacg ggactttcct acttggcagt acatctacgt attagtcatc gctattacca    3900
tggtgatgcg gttttggcag tacatcaatg ggcgtggata gcggtttgac tcacggggat    3960
ttccaagtct ccaccccatt gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg    4020
actttccaaa atgtcgtaac aactccgccc cattgacgca aatgggcggt aggcgtgtac    4080
ggtgggaggt ctatataagc agagctcgtt tagtgaaccg tcagatcgcc tggagacgcc    4140
atccacgctg ttttgacctc catagaagac accgggaccg atccagcctc ccctcgaagc    4200
tcgactctag gggctcgaga tctgcgatct aagtaagctt gcatgcctgc aggtcggccg    4260
ccacgaccgg tgccgccacc atcccctgac ccacgcccct gacccctcac aaggagacga    4320
ccttccatga ccgagtacaa gcccacggtg cgcctcgcca cccgcgacga cgtcccccgg    4380
gccgtacgca ccctcgccgc cgcgttcgcc gactaccccg ccacgcgcca caccgtcgac    4440
ccggaccgcc acatcgagcg ggtcaccgag ctgcaagaac tcttcctcac gcgcgtcggg    4500
ctcgacatcg gcaaggtgtg ggtcgcggac gacggcgccg cggtggcggt ctggaccacg    4560
ccggagagcg tcgaagcggg ggcggtgttc gccgagatcg cccgcgcat ggccgagttg    4620
agcggttccc ggctggccgc gcagcaacag atggaaggcc tcctggcgcc gcaccggccc    4680
aaggagcccg cgtggttcct ggccaccgtc ggcgtctcgc ccgaccacca gggcaagggt    4740
ctgggcagcg ccgtcgtgct ccccggagtg gaggcggccg agcgcgccgg ggtgcccgcc    4800
ttcctggaga cctccgcgcc ccgcaacctc cccttctacg agcggctcgg cttcaccgtc    4860
accgccgacg tcgaggtgcc cgaaggaccg cgcacctggt gcatgacccg caagcccggt    4920
gcctgacgcc cgccccacga cccgcagcgc ccgaccgaaa ggagcgcacg accccatggc    4980
tccgaccgaa gccgacccgg cggccccgc cgaccccgca cccgcccccg aggcccaccg    5040
act                                                                  5043
```

```
<210>   6
<211>   5041
<212>   DNA
<213>   Plasmid pCMV-EGFP-attB

<400>   6
ctagagtcgg ggcggccggc cgcttcgagc agacatgata agatacattg atgagtttgg     60
acaaaccaca actagaatgc agtgaaaaaa atgctttatt tgtgaaattt gtgatgctat    120
tgctttattt gtaaccatta taagctgcaa taaacaagtt aacaacaaca attgcattca    180
ttttatgttt caggttcagg gggaggtgtg ggaggttttt taaagcaagt aaaacctcta    240
caaatgtggt aaaatcgata aggatcaatt cggcttcagg taccgtcgac gatgtaggtc    300
acggtctcga gccgcggtg cgggtgccag ggcgtgccct gggctcccc gggcgcgtac    360
tccacctcac ccatctggtc catcatgatg aacgggtcga ggtggcggta gttgatcccg    420
gcgaacgcgc ggcgcaccgg gaagccctcg ccctcgaaac cgctgggcgc ggtggtcacg    480
gtgagcacgg gacgtgcgac ggcgtcggcg ggtgcggata cgcggggcag cgtcagcggg    540
ttctcgacgg tcacggcggg catgtcgaca gccgaattga tccgtcgacc gatgcccttg    600
```
                              Page 15

Integrase.ST25

agagccttca	acccagtcag	ctccttccgg	tgggcgcggg	gcatgactat	cgtcgccgca	660
cttatgactg	tcttctttat	catgcaactc	gtaggacagg	tgccggcagc	gctcttccgc	720
ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	cggcgagcgg	tatcagctca	780
ctcaaaggcg	gtaatacggt	tatccacaga	atcaggggat	aacgcaggaa	agaacatgtg	840
agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	gcgttgctgg	cgtttttcca	900
taggctccgc	cccctgacg	agcatcacaa	aaatcgacgc	tcaagtcaga	ggtggcgaaa	960
cccgacagga	ctataaagat	accaggcgtt	tcccctgga	agctccctcg	tgcgctctcc	1020
tgttccgacc	ctgccgctta	ccggatacct	gtccgccttt	ctcccttcgg	gaagcgtggc	1080
gctttctcaa	tgctcacgct	gtaggtatct	cagttcggtg	taggtcgttc	gctccaagct	1140
gggctgtgtg	cacgaacccc	ccgttcagcc	cgaccgctgc	gccttatccg	gtaactatcg	1200
tcttgagtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcca	ctggtaacag	1260
gattagcaga	gcgaggtatg	taggcggtgc	tacagagttc	ttgaagtggt	ggcctaacta	1320
cggctacact	agaaggacag	tatttggtat	ctgcgctctg	ctgaagccag	ttaccttcgg	1380
aaaaagagtt	ggtagctctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggtttttt	1440
tgtttgcaag	cagcagatta	cgcgcagaaa	aaaaggatct	caagaagatc	ctttgatctt	1500
ttctacgggg	tctgacgctc	agtggaacga	aaactcacgt	taagggattt	tggtcatgag	1560
attatcaaaa	aggatcttca	cctagatcct	tttaaattaa	aaatgaagtt	ttaaatcaat	1620
ctaaagtata	tatgagtaaa	cttggtctga	cagttaccaa	tgcttaatca	gtgaggcacc	1680
tatctcagcg	atctgtctat	ttcgttcatc	catagttgcc	tgactccccg	tcgtgtagat	1740
aactacgata	cgggagggct	taccatctgg	ccccagtgct	gcaatgatac	cgcgagaccc	1800
acgctcaccg	gctccagatt	tatcagcaat	aaaccagcca	gccggaaggg	ccgagcgcag	1860
aagtggtcct	gcaactttat	ccgcctccat	ccagtctatt	aattgttgcc	gggaagctag	1920
agtaagtagt	tcgccagtta	atagtttgcg	caacgttgtt	gccattgcta	caggcatcgt	1980
ggtgtcacgc	tcgtcgtttg	gtatggcttc	attcagctcc	ggttcccaac	gatcaaggcg	2040
agttacatga	tcccccatgt	tgtgcaaaaa	agcggttagc	tccttcggtc	ctccgatcgt	2100
tgtcagaagt	aagttggccg	cagtgttatc	actcatggtt	atggcagcac	tgcataattc	2160
tcttactgtc	atgccatccg	taagatgctt	ttctgtgact	ggtgagtact	caaccaagtc	2220
attctgagaa	tagtgtatgc	ggcgaccgag	ttgctcttgc	ccggcgtcaa	tacgggataa	2280
taccgcgcca	catagcagaa	ctttaaaagt	gctcatcatt	ggaaaacgtt	cttcggggcg	2340
aaaactctca	aggatcttac	cgctgttgag	atccagttcg	atgtaaccca	ctcgtgcacc	2400
caactgatct	tcagcatctt	ttactttcac	cagcgtttct	gggtgagcaa	aaacaggaag	2460
gcaaaatgcc	gcaaaaaagg	gaataagggc	gacacggaaa	tgttgaatac	tcatactctt	2520
ccttttttcaa	tattattgaa	gcatttatca	gggttattgt	ctcatgagcg	gatacatatt	2580
tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgcgc	acatttcccc	gaaaagtgcc	2640

Integrase.ST25

acctgacgcg	ccctgtagcg	gcgcattaag	cgcggcgggt	gtggtggtta	cgcgcagcgt	2700
gaccgctaca	cttgccagcg	ccctagcgcc	cgctcctttc	gctttcttcc	cttcctttct	2760
cgccacgttc	gccggctttc	cccgtcaagc	tctaaatcgg	gggctccctt	tagggttccg	2820
atttagtgct	ttacggcacc	tcgaccccaa	aaaacttgat	tagggtgatg	gttcacgtag	2880
tgggccatcg	ccctgataga	cggttttttcg	ccctttgacg	ttggagtcca	cgttctttaa	2940
tagtggactc	ttgttccaaa	ctggaacaac	actcaaccct	atctcggtct	attcttttga	3000
tttataaggg	attttgccga	tttcggccta	ttggttaaaa	aatgagctga	tttaacaaaa	3060
atttaacgcg	aattttaaca	aaatattaac	gtttacaatt	tcccattcgc	cattcaggct	3120
gcgcaactgt	tgggaagggc	gatcggtgcg	ggcctcttcg	ctattacgcc	agcccaagct	3180
accatgataa	gtaagtaata	ttaaggtacg	ggaggtactt	ggagcggccg	caataaaata	3240
tctttatttt	cattacatct	gtgtgttggt	tttttgtgtg	aatcgatagt	actaacatac	3300
gctctccatc	aaaacaaaac	gaaacaaaac	aaactagcaa	aataggctgt	ccccagtgca	3360
agtgcaggtg	ccagaacatt	tctctatcga	taggtaccga	gctcttacgc	gtgctagccc	3420
tcgagcagga	tctatacatt	gaatcaatat	tggcaattag	ccatattagt	cattggttat	3480
atagcataaa	tcaatattgg	ctattggcca	ttgcatacgt	tgtatctata	tcataatatg	3540
tacatttata	ttggctcatg	tccaatatga	ccgccatgtt	gacattgatt	attgactagt	3600
tattaatagt	aatcaattac	ggggtcatta	gttcatagcc	catatatgga	gttccgcgtt	3660
acataactta	cggtaaatgg	cccgcctggc	tgaccgccca	acgacccccg	cccattgacg	3720
tcaataatga	cgtatgttcc	catagtaacg	ccaatagggа	ctttccattg	acgtcaatgg	3780
gtggagtatt	tacggtaaac	tgcccacttg	gcagtacatc	aagtgtatca	tatgccaagt	3840
ccgcccccta	ttgacgtcaa	tgacggtaaa	tggcccgcct	ggcattatgc	ccagtacatg	3900
accttacggg	actttcctac	ttggcagtac	atctacgtat	tagtcatcgc	tattaccatg	3960
gtgatgcggt	tttggcagta	catcaatggg	cgtggatagc	ggtttgactc	acggggattt	4020
ccaagtctcc	accccattga	cgtcaatggg	agtttgtttt	ggcaccaaaa	tcaacgggac	4080
tttccaaaat	gtcgtaacaa	ctccgcccca	ttgacgcaaa	tgggcggtag	gcgtgtacgg	4140
tgggaggtct	atataagcag	agctcgttta	gtgaaccgtc	agatcgcctg	gagacgccat	4200
ccacgctgtt	ttgacctcca	tagaagacac	cgggaccgat	ccagcctccc	ctcgaagctc	4260
gactctaggg	gctcgagatc	cccgggtacc	ggtcgccacc	atggtgagca	agggcgagga	4320
gctgttcacc	ggggtggtgc	ccatcctggt	cgagctggac	ggcgacgtaa	acggccacaa	4380
gttcagcgtg	tccggcgagg	gcgagggcga	tgccacctac	ggcaagctga	ccctgaagtt	4440
catctgcacc	accggcaagc	tgcccgtgcc	ctggcccacc	ctcgtgacca	ccctgaccta	4500
cggcgtgcag	tgcttcagcc	gctaccccga	ccacatgaag	cagcacgact	tcttcaagtc	4560
cgccatgccc	gaaggctacg	tccaggagcg	caccatcttc	ttcaaggacg	acggcaacta	4620
caagacccgc	gccgaggtga	agttcgaggg	cgacaccctg	gtgaaccgca	tcgagctgaa	4680

Page 17

Integrase.ST25

```
gggcatcgac ttcaaggagg acggcaacat cctgggggcac aagctggagt acaactacaa    4740
cagccacaac gtctatatca tggccgacaa gcagaagaac ggcatcaagg tgaacttcaa    4800
gatccgccac aacatcgagg acggcagcgt gcagctcgcc gaccactacc agcagaacac    4860
ccccatcggc gacggccccg tgctgctgcc cgacaaccac tacctgagca cccagtccgc    4920
cctgagcaaa gaccccaacg agaagcgcga tcacatggtc ctgctggagt tcgtgaccgc    4980
cgccgggatc actctcggca tggacgagct gtacaagtaa agcggccgct cgagcatgca    5040
t                                                                    5041
```

```
<210>   7
<211>   18116
<212>   DNA
<213>   Plasmid p12.0lys-LSPIPNMM-CMV-pur-attB

<400>   7
gggctgcagg aattcgattg ccgccttctt tgatattcac tctgttgtat ttcatctctt      60
cttgccgatg aaaggatata acagtctgta taacagtctg tgaggaaata cttggtattt     120
cttctgatca gtgttttttat aagtaatgtt gaatattgga taaggctgtg tgtcctttgt     180
cttgggagac aaagcccaca gcaggtggtg gttggggtgg tggcagctca gtgacaggag     240
aggttttttt gcctgttttt ttttttttttt tttttttaa gtaaggtgtt cttttttctt     300
agtaaatttt ctactggact gtatgttttg acaggtcaga aacatttctt caaaagaaga     360
accttttgga aactgtacag ccctttttctt tcattccctt tttgctttct gtgccaatgc     420
ctttggttct gattgcatta tggaaaacgt tgatcggaac ttgaggtttt tatttatagt     480
gtggcttgaa agcttggata gctgttgtta cacgagatac cttattaagt ttaggccagc     540
ttgatgcttt attttttccc tttgaagtag tgagcgttct ctggtttttt tcctttgaaa     600
ctggtgaggc ttagatttttt ctaatgggat ttttttacctg atgatctagt tgcataccca     660
aatgcttgta aatgttttcc tagttaacat gttgataact tcggatttac atgttgtata     720
tacttgtcat ctgtgtttct agtaaaaata tatggcattt atagaaatac gtaattcctg     780
atttcctttt tttttatctc tatgctctgt gtgtacaggt caaacagact tcactcctat     840
ttttatttat agaattttat atgcagtctg tcgttggttc ttgtgttgta aggatacagc     900
cttaaatttc ctagagcgat gctcagtaag gcgggttgtc acatgggttc aaatgtaaaa     960
cgggcacgtt tggctgctgc cttcccgaga tccaggacac taaactgctt ctgcactgag    1020
gtataaatcg cttcagatcc cagggaagtg cagatccacg tgcatattct aaagaagaa    1080
tgaatacttt ctaaaatatt ttggcatagg aagcaagctg catggatttg tttgggactt    1140
aaattatttt ggtaacggag tgcataggtt ttaaacacag ttgcagcatg ctaacgagtc    1200
acagcgttta tgcagaagtg atgcctggat gcctgttgca gctgtttacg gcactgcctt    1260
gcagtgagca ttgcagatag gggtggggtg ctttgtgtcg tgttcccaca cgctgccaca    1320
cagccacctc ccggaacaca tctcacctgc tgggtacttt tcaaaccatc ttagcagtag    1380
```

Page 18

Integrase.ST25

```
tagatgagtt actatgaaac agagaagttc ctcagttgga tattctcatg ggatgtcttt   1440

tttcccatgt tgggcaaagt atgataaagc atctctattt gtaaattatg cacttgttag   1500

ttcctgaatc ctttctatag caccacttat tgcagcaggt gtaggctctg gtgtggcctg   1560

tgtctgtgct tcaatctttt aaagcttctt tggaaataca ctgacttgat tgaagtctct   1620

tgaagatagt aaacagtact taccttttgat cccaatgaaa tcgagcattt cagttgtaaa   1680

agaattccgc ctattcatac catgtaatgt aattttacac ccccagtgct gacactttgg   1740

aatatattca agtaatagac tttggcctca ccctcttgtg tactgtattt tgtaatagaa   1800

aatattttaa actgtgcata tgattattac attatgaaag agacattctg ctgatcttca   1860

aatgtaagaa aatgaggagt gcgtgtgctt ttataaatac aagtgattgc aaattagtgc   1920

aggtgtcctt aaaaaaaaaa aaaaaaagta atataaaaag gaccaggtgt tttacaagtg   1980

aaatacattc ctatttggta aacagttaca tttttatgaa gattaccagc gctgctgact   2040

ttctaaacat aaggctgtat tgtcttcctg taccattgca tttcctcatt cccaatttgc   2100

acaaggatgt ctgggtaaac tattcaagaa atggctttga aatacagcat gggagcttgt   2160

ctgagttgga atgcagagtt gcactgcaaa atgtcaggaa atggatgtct ctcagaatgc   2220

ccaactccaa aggattttat atgtgtatat agtaagcagt ttcctgattc cagcaggcca   2280

aagagtctgc tgaatgttgt gttgccggag acctgtattt ctcaacaagg taagatggta   2340

tcctagcaac tgcggatttt aatacatttt cagcagaagt acttagttaa tctctacctt   2400

tagggatcgt ttcatcattt ttagatgtta tacttgaaat actgcataac ttttagcttt   2460

catgggttcc ttttttttcag cctttaggag actgttaagc aatttgctgt ccaacttttg   2520

tgttggtctt aaactgcaat agtagtttac cttgtattga agaaataaag accattttta   2580

tattaaaaaa tacttttgtc tgtcttcatt ttgacttgtc tgatatcctt gcagtgccca   2640

ttatgtcagt tctgtcagat attcagacat caaaacttaa cgtgagctca gtggagttac   2700

agctgcggtt ttgatgctgt tattatttct gaaactagaa atgatgttgt cttcatctgc   2760

tcatcaaaca cttcatgcag agtgtaaggc tagtgagaaa tgcatacatt tattgatact   2820

tttttaaagt caacttttta tcagattttt ttttcatttg gaaatatatt gttttctaga   2880

ctgcatagct tctgaatctg aaatgcagtc tgattggcat gaagaagcac agcactcttc   2940

atcttactta aacttcattt tggaatgaag gaagttaagc aagggcacag gtccatgaaa   3000

tagagacagt gcgctcagga gaaagtgaac ctggatttct ttggctagtg ttctaaatct   3060

gtagtgagga aagtaacacc cgattccttg aaagggctcc agctttaatg cttccaaatt   3120

gaaggtggca ggcaacttgg ccactggtta tttactgcat tatgtctcag tttcgcagct   3180

aacctggctt ctccactatt gagcatggac tatagcctgg cttcagaggc caggtgaagg   3240

ttgggatggg tggaaggagt gctgggctgt ggctggggGG actgtgggga ctccaagctg   3300

agcttggggt gggcagcaca gggaaaagtg tgggtaacta ttttttaagta ctgtgttgca   3360

aacgtctcat ctgcaaatac gtagggtgtg tactctcgaa gattaacagt gtgggttcag   3420
```

Page 19

Integrase.ST25
```
taatatatgg atgaattcac agtggaagca ttcaagggta gatcatctaa cgacaccaga      3480
tcatcaagct atgattggaa gcggtatcag aagagcgagg aaggtaagca gtcttcatat      3540
gttttccctc cacgtaaagc agtctgggaa agtagcaccc cttgagcaga gacaaggaaa      3600
taattcagga gcatgtgcta ggagaacttt cttgctgaat tctacttgca agagctttga      3660
tgcctggctt ctggtgcctt ctgcagcacc tgcaaggccc agagcctgtg gtgagctgga      3720
gggaaagatt ctgctcaagt ccaagcttca gcaggtcatt gtctttgctt cttcccccag      3780
cactgtgcag cagagtggaa ctgatgtcga agcctcctgt ccactacctg ttgctgcagg      3840
cagactgctc tcagaaaaag agagctaact ctatgccata gtctgaaggt aaaatgggtt      3900
ttaaaaaaga aaacacaaag gcaaaaccgg ctgccccatg agaagaaagc agtggtaaac      3960
atggtagaaa aggtgcagaa gcccccaggc agtgtgacag gcccctcctg ccacctagag      4020
gcgggaacaa gcttccctgc ctagggctct gcccgcgaag tgcgtgtttc tttggtgggt      4080
tttgtttggc gtttggtttt gagatttaga cacaaggaa gcctgaaagg aggtgttggg      4140
cactattttg gtttgtaaag cctgtacttc aaatatatat tttgtgaggg agtgtagcga      4200
attggccaat ttaaaataaa gttgcaagag attgaaggct gagtagttga gagggtaaca      4260
cgtttaatga gatcttctga aactactgct tctaaacact tgtttgagtg gtgagacctt      4320
ggataggtga gtgctcttgt tacatgtctg atgcacttgc ttgtcctttt ccatccacat      4380
ccatgcattc cacatccacg catttgtcac ttatcccata tctgtcatat ctgacatacc      4440
tgtctcttcg tcacttggtc agaagaaaca gatgtgataa tccccagccg ccccaagttt      4500
gagaagatgg cagttgcttc tttccctttt tcctgctaag taaggatttt ctcctggctt      4560
tgacacctca cgaaatagtc ttcctgcctt acattctggg cattatttca aatatctttg      4620
gagtgcgctg ctctcaagtt tgtgtcttcc tactcttaga gtgaatgctc ttagagtgaa      4680
agagaaggaa gagaagatgt tggccgcagt tctctgatga acacacctct gaataatggc      4740
caaaggtggg tgggtttctc tgaggaacgg gcagcgtttg cctctgaaag caaggagctc      4800
tgcggagttg cagttatttt gcaactgatg gtggaactgg tgcttaaagc agattcccta      4860
ggttccctgc tacttctttt ccttcttggc agtcagttta tttctgacag acaaacagcc      4920
accccactg caggcttaga aagtatgtgg ctctgcctgg gtgtgttaca gctctgccct      4980
ggtgaaaggg gattaaaacg gcaccattc atcccaaaca ggatcctcat tcatggatca      5040
agctgtaagg aacttgggct ccaacctcaa aacattaatt ggagtacgaa tgtaattaaa      5100
actgcattct cgcattccta agtcatttag tctggactct gcagcatgta ggtcggcagc      5160
tcccactttc tcaaagacca ctgatggagg agtagtaaaa atggagaccg attcagaaca      5220
accaacggag tgttgccgaa gaaactgatg gaaataatgc atgaattgtg tggtggacat      5280
ttttttaaa tacataaact acttcaaatg aggtcggaga aggtcagtgt tttattagca      5340
gccataaaac caggtgagcg agtaccattt ttctctacaa gaaaaacgat tctgagctct      5400
gcgtaagtat aagttctcca tagcggctga agctcccccc tggctgcctg ccatctcagc      5460
```
Page 20

Integrase.ST25

```
tggagtgcag tgccatttcc ttggggtttc tctcacagca gtaatgggac aatacttcac    5520
aaaaattctt tcttttcctg tcatgtggga tccctactgt gccctcctgg ttttacgtta    5580
ccccctgact gttccattca gcggtttgga aagagaaaaa gaatttggaa ataaaacatg    5640
tctacgttat cacctcctcc agcattttgg tttttaatta tgtcaataac tggcttagat    5700
ttggaaatga gaggggggttg ggtgtattac cgaggaacaa aggaaggctt atataaactc    5760
aagtctttta tttagagaac tggcaagctg tcaaaaacaa aaaggcctta ccaccaaatt    5820
aagtgaatag ccgctatagc cagcagggcc agcacgaggg atggtgcact gctggcacta    5880
tgccacggcc tgcttgtgac tctgagagca actgctttgg aaatgacagc acttggtgca    5940
atttcctttg tttcagaatg cgtagagcgt gtgcttggcg acagtttttc tagttaggcc    6000
acttcttttt tccttctctc ctcattctcc taagcatgtc tccatgctgg taatcccagt    6060
caagtgaacg ttcaaacaat gaatccatca ctgtaggatt ctcgtggtga tcaaatcttt    6120
gtgtgaggtc tataaaatat ggaagcttat ttatttttcg ttcttccata tcagtcttct    6180
ctatgacaat tcacatccac cacagcaaat taaaggtgaa ggaggctggt gggatgaaga    6240
gggtcttcta gctttacgtt cttccttgca aggccacagg aaaatgctga gagctgtaga    6300
atacagcctg gggtaagaag ttcagtctcc tgctgggaca gctaaccgca tcttataacc    6360
ccttctgaga ctcatcttag gaccaaatag ggtctatctg gggttttttgt tcctgctgtt    6420
cctcctggaa ggctatctca ctatttcact gctcccacgg ttacaaacca aagatacagc    6480
ctgaattttt tctaggccac attacataaa tttgacctgg taccaatatt gttctctata    6540
tagttatttc cttccccact gtgtttaacc ccttaaggca ttcagaacaa ctagaatcat    6600
agaatggttt ggattggaag gggccttaaa catcatccat ttccaaccct ctgccatggg    6660
ctgcttgcca cccactggct caggctgccc agggccccat ccagcctggc cttgagcacc    6720
tccagggatg gggcacccac agcttctctg ggcagcctgt gccaacacct caccactctc    6780
tgggtaaaga attctctttt aacatctaat ctaaatctct tctcttttag tttaaagcca    6840
ttcctctttt tcccgttgct atctgtccaa gaaatgtgta ttggtctccc tcctgcttat    6900
aagcaggaag tactggaagg ctgcagtgag gtctccccac agccttctct tctccaggct    6960
gaacaagccc agctccttca gcctgtcttc gtaggagatc atcttagtgg ccctcctctg    7020
gacccattcc aacagttcca cggctttctt gtggagcccc aggtctggat gcagtacttc    7080
agatggggcc ttacaaaggc agagcagatg gggacaatcg cttacccctc cctgctggct    7140
gccctgtttt tgatgcagcc cagggtactg ttggcctttc aggctcccag accccttgct    7200
gatttgtgtc aagcttttca tccaccagaa cccacgcttc ctggttaata cttctgccct    7260
cacttctgta agcttgtttc aggagacttc cattctttag gacagactgt gttacaccta    7320
cctgccctat tcttgcatat atacatttca gttcatgttt cctgtaacag gacagaatat    7380
gtattcctct aacaaaaata catgcagaat tcctagtgcc atctcagtag ggttttcatg    7440
gcagtattag cacatagtca atttgctgca agtaccttcc aagctgcggc ctcccataaa    7500
```

Page 21

Integrase.ST25

```
tcctgtatttt  gggatcagtt  accttttggg  gtaagctttt  gtatctgcag  agaccctggg    7560
ggttctgatg  tgcttcagct  ctgctctgtt  ctgactgcac  cattttctag  atcacccagt    7620
tgttcctgta  caacttcctt  gtcctccatc  ctttcccagc  ttgtatcttt  gacaaataca    7680
ggcctatttt  tgtgtttgct  tcagcagcca  tttaattctt  cagtgtcatc  ttgttctgtt    7740
gatgccactg  gaacaggatt  ttcagcagtc  ttgcaaagaa  catctagctg  aaaactttct    7800
gccattcaat  attcttacca  gttcttcttg  tttgaggtga  gccataaatt  actagaactt    7860
cgtcactgac  aagtttatgc  attttattac  ttctattatg  tacttacttt  gacataacac    7920
agacacgcac  atattttgct  gggatttcca  cagtgtctct  gtgtccttca  catggtttta    7980
ctgtcatact  tccgttataa  ccttggcaat  ctgcccagct  gcccatcaca  agaaaagaga    8040
ttcctttttt  attacttctc  ttcagccaat  aaacaaaatg  tgagaagccc  aaacaagaac    8100
ttgtggggca  ggctgccatc  aagggagaga  cagctgaagg  gttgtgtagc  tcaatagaat    8160
taagaaataa  taaagctgtg  tcagacagtt  ttgcctgatt  tatacaggca  cgccccaagc    8220
cagagaggct  gtctgccaag  gccaccttgc  agtccttggt  ttgtaagata  agtcataggt    8280
aacttttctg  gtgaattgcg  tggagaatca  tgatggcagt  tcttgctgtt  tactatggta    8340
agatgctaaa  ataggagaca  gcaaagtaac  acttgctgct  gtaggtgctc  tgctatccag    8400
acagcgatgg  cactcgcaca  ccaagatgag  ggatgctccc  agctgacgga  tgctggggca    8460
gtaacagtgg  gtcccatgct  gcctgctcat  tagcatcacc  tcagccctca  ccagcccatc    8520
agaaggatca  tcccaagctg  aggaaagttg  ctcatcttct  tcacatcatc  aaacctttgg    8580
cctgactgat  gcctcccgga  tgcttaaatg  tggtcactga  catctttatt  tttctatgat    8640
ttcaagtcag  aacctccgga  tcaggaggga  acacatagtg  ggaatgtacc  ctcagctcca    8700
aggccagatc  ttccttcaat  gatcatgcat  gctacttagg  aaggtgtgtg  tgtgtgaatg    8760
tagaattgcc  tttgttattt  tttcttcctg  ctgtcaggaa  cattttgaat  accagagaaa    8820
aagaaaagtg  ctcttcttgg  catgggagga  gttgtcacac  ttgcaaaata  aaggatgcag    8880
tcccaaatgt  tcataatctc  agggtctgaa  ggaggatcag  aaactgtgta  tacaatttca    8940
ggcttctctg  aatgcagctt  ttgaaagctg  ttcctggccg  aggcagtact  agtcagaacc    9000
ctcggaaaca  ggaacaaatg  tcttcaggt  gcagcaggag  gaaacacctt  gcccatcatg    9060
aaagtgaata  accactgccg  ctgaaggaat  ccagctcctg  tttgagcagg  tgctgcacac    9120
tcccacactg  aaacaacagt  tcattttat  aggacttcca  ggaaggatct  tcttcttaag    9180
cttcttaatt  atggtacatc  tccagttggc  agatgactat  gactactgac  aggagaatga    9240
ggaactagct  gggaatattt  ctgtttgacc  accatggagt  cacccatttc  tttactggta    9300
tttggaaata  ataattctga  attgcaaagc  aggagttagc  gaagatcttc  atttcttcca    9360
tgttggtgac  agcacagttc  tggctatgaa  agtctgctta  caaggaagag  gataaaaatc    9420
atagggataa  taaatctaag  tttgaagaca  atgaggtttt  agctgcattt  gacatgaaga    9480
aattgagacc  tctactggat  agctatggta  tttacgtgtc  tttttgctta  gttacttatt    9540
```

Page 22

Integrase.ST25

```
gaccccagct gaggtcaagt atgaactcag gtctctcggg ctactggcat ggattgatta   9600
catacaactg taattttagc agtgatttag ggtttatgag tacttttgca gtaaatcata   9660
gggttagtaa tgttaatctc agggaaaaaa aaaaaaagcc aaccctgaca gacatcccag   9720
ctcaggtgga aatcaaggat cacagctcag tgcggtccca gagaacacag ggactcttct   9780
cttaggacct ttatgtacag ggcctcaaga taactgatgt tagtcagaag actttccatt   9840
ctggccacag ttcagctgag gcaatcctgg aattttctct ccgctgcaca gttccagtca   9900
tcccagtttg tacagttctg gcactttttg ggtcaggccg tgatccaagg agcagaagtt   9960
ccagctatgg tcagggagtg cctgaccgtc ccaactcact gcactcaaac aaaggcgaaa  10020
ccacaagagt ggcttttgtt gaaattgcag tgtggcccag aggggctgca ccagtactgg  10080
attgaccacg aggcaacatt aatcctcagc aagtgcaatt tgcagccatt aaattgaact  10140
aactgatact acaatgcaat cagtatcaac aagtggtttg gcttggaaga tggagtctag  10200
gggctctaca ggagtagcta ctctctaatg gagttgcatt ttgaagcagg acactgtgaa  10260
aagctggcct cctaaagagg ctgctaaaca ttagggtcaa ttttccagtg cactttctga  10320
agtgtctgca gttccccatg caaagctgcc caaacatagc acttccaatt gaatacaatt  10380
atatgcaggc gtactgcttc ttgccagcac tgtccttctc aaatgaactc aacaaacaat  10440
ttcaaagtct agtagaaagt aacaagcttt gaatgtcatt aaaaagtata tctgctttca  10500
gtagttcagc ttatttatgc ccactagaaa catcttgtac aagctgaaca ctggggctcc  10560
agattagtgg taaaacctac tttatacaat catagaatca tagaatggcc tgggttggaa  10620
gggaccccaa ggatcatgaa gatccaacac ccccgccaca ggcagggcca ccaacctcca  10680
gatctggtac tagaccaggc agcccagggc tccatccaac ctggccatga acacctccag  10740
ggatggagca tccacaacct ctctgggcag cctgtgccag cacctcacca ccctctctgt  10800
gaagaacttt tccctgacat ccaatctaag ccttccctcc ttgaggttag atccactccc  10860
ccttgtgcta tcactgtcta ctcttgtaaa aagttgattc tcctcctttt tggaaggttg  10920
caatgaggtc tccttgcagc cttcttctct tctgcaggat gaacaagccc agctccctca  10980
gcctgtcttt ataggagagg tgctccagcc ctctgatcat ctttgtggcc ctcctctgga  11040
cccgctccaa gagctccaca tctttcctgt actgggggcc ccaggcctga atgcagtact  11100
ccagatgggg cctcaaaaga gcagagtaaa gagggacaat caccttcctc accctgctgg  11160
ccagccctct tctgatggag ccctggatac aactggcttt ctgagctgca acttctcctt  11220
atcagttcca ctattaaaac aggaacaata caacaggtgc tgatggccag tgcagagttt  11280
ttcacacttc ttcatttcgg tagatcttag atgaggaacg ttgaagttgt gcttctgcgt  11340
gtgcttcttc ctcctcaaat actcctgcct gatacctcac cccacctgcc actgaatggc  11400
tccatggccc cctgcagcca gggccctgat gaacccggca ctgcttcaga tgctgtttaa  11460
tagcacagta tgaccaagtt gcacctatga atacacaaac aatgtgttgc atccttcagc  11520
acttgagaag aagagccaaa tttgcattgt caggaaatgg tttagtaatt ctgccaatta  11580
```

Integrase.ST25

```
aaacttgttt atctaccatg gctgttttta tggctgttag tagtggtaca ctgatgatga   11640
acaatggcta tgcagtaaaa tcaagactgt agatattgca acagactata aaattcctct   11700
gtggcttagc caatgtggta cttcccacat tgtataagaa atttggcaag tttagagcaa   11760
tgtttgaagt gttgggaaat ttctgtatac tcaagagggc gttttttgaca actgtagaac   11820
agaggaatca aaaggggtg ggaggaagtt aaaagaagag gcaggtgcaa gagagcttgc   11880
agtcccgctg tgtgtacgac actggcaaca tgaggtcttt gctaatcttg gtgctttgct   11940
tcctgcccct ggctgcctta gggtgcgatc tgcctcagac ccacagcctg ggcagcagga   12000
ggaccctgat gctgctggct cagatgagga gaatcagcct gtttagctgc ctgaaggata   12060
ggcacgattt tggctttcct caagaggagt ttggcaacca gtttcagaag gctgagacca   12120
tccctgtgct gcacgagatg atccagcaga tctttaacct gtttagcacc aaggatagca   12180
gcgctgcttg ggatgagacc ctgctggata agttttacac cgagctgtac cagcagctga   12240
acgatctgga ggcttgcgtg atccagggcg tgggcgtgac cgagacccct ctgatgaagg   12300
aggatagcat cctggctgtg aggaagtact ttcagaggat caccctgtac ctgaaggaga   12360
agaagtacag cccctgcgct tgggaagtcg tgagggctga gatcatgagg agctttagcc   12420
tgagcaccaa cctgcaagag agcttgaggt ctaaggagta aaaagtctag agtcggggcg   12480
gccggccgct tcgagcagac atgataagat acattgatga gtttggacaa accacaacta   12540
gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct ttatttgtaa   12600
ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt atgtttcagg   12660
ttcagggga ggtgtgggag gttttttaaa gcaagtaaaa cctctacaaa tgtggtaaaa   12720
tcgataagga tccgtcgacc gatgcccttg agagccttca acccagtcag ctccttccgg   12780
tgggcgcggg gcatgactat cgtcgccgca cttatgactg tcttctttat catgcaactc   12840
gtaggacagg tgccggcagc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt   12900
cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga   12960
atcaggggat aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg   13020
taaaaaggcc gcgttgctgg cgttttttcca taggctccgc cccctgacg agcatcacaa   13080
aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt   13140
tcccctgga agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct   13200
gtccgccttt ctcccttcgg gaagcgtggc gctttctcaa tgctcacgct gtaggtatct   13260
cagttcggtg taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc   13320
cgaccgctgc gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt   13380
atcgccactg gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc   13440
tacagagttc ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat   13500
ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa   13560
acaaaccacc gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa   13620
```

Page 24

Integrase.ST25

```
aaaaggatct caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga   13680
aaactcacgt taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct   13740
tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga   13800
cagttaccaa tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc   13860
catagttgcc tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg   13920
ccccagtgct gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat   13980
aaaccagcca gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat   14040
ccagtctatt aattgttgcc gggaagctag agtaagtagt tcgccagtta atagtttgcg   14100
caacgttgtt gccattgcta caggcatcgt ggtgtcacgc tcgtcgtttg gtatggcttc   14160
attcagctcc ggttcccaac gatcaaggcg agttacatga tcccccatgt tgtgcaaaaa   14220
agcggttagc tccttcggtc ctccgatcgt tgtcagaagt aagttggccg cagtgttatc   14280
actcatggtt atggcagcac tgcataattc tcttactgtc atgccatccg taagatgctt   14340
ttctgtgact ggtgagtact caaccaagtc attctgagaa tagtgtatgc ggcgaccgag   14400
ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca catagcagaa ctttaaaagt   14460
gctcatcatt ggaaaacgtt cttcggggcg aaaactctca aggatcttac cgctgttgag   14520
atccagttcg atgtaaccca ctcgtgcacc caactgatct tcagcatctt ttactttcac   14580
cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc gcaaaaaagg gaataagggc   14640
gacacggaaa tgttgaatac tcatactctt ccttttttcaa tattattgaa gcatttatca   14700
gggttattgt ctcatgagcg gatacatatt tgaatgtatt tagaaaaata aacaaatagg   14760
ggttccgcgc acatttcccc gaaaagtgcc acctgacgcg ccctgtagcg gcgcattaag   14820
cgcggcgggt gtggtggtta cgcgcagcgt gaccgctaca cttgccagcg ccctagcgcc   14880
cgctcctttc gctttcttcc cttcctttct cgccacgttc gccggctttc cccgtcaagc   14940
tctaaatcgg gggctccctt tagggttccg atttagtgct ttacggcacc tcgaccccaa   15000
aaaacttgat tagggtgatg gttcacgtag tgggccatcg ccctgataga cggttttttcg   15060
cccttttgacg ttggagtcca cgttctttaa tagtggactc ttgttccaaa ctggaacaac   15120
actcaaccct atctcggtct attcttttga tttataaggg attttgccga tttcggccta   15180
ttggttaaaa aatgagctga tttaacaaaa atttaacgcg aattttaaca aaatattaac   15240
gtttacaatt tcccattcgc cattcaggct gcgcaactgt tgggaagggc gatcggtgcg   15300
ggcctcttcg ctattacgcc agcccaagct accatgataa gtaagtaata ttaaggtacg   15360
ggaggtactt ggagcggccg ctctagaact agtggatccc ccggccgcaa taaaatatct   15420
ttattttcat tacatctgtg tgttggtttt ttgtgtgaat cgatagtact aacatacgct   15480
ctccatcaaa acaaaacgaa acaaaacaaa ctagcaaaat aggctgtccc cagtgcaagt   15540
gcaggtgcca gaacatttct ctatcgatag gtaccgagct cttacgcgtg ctagccctcg   15600
agcaggatct atacattgaa tcaatattgg caattagcca tattagtcat tggttatata   15660
```

Page 25

Integrase.ST25

```
gcataaatca atattggcta ttggccattg catacgttgt atctatatca taatatgtac   15720
atttatattg gctcatgtcc aatatgaccg ccatgttgac attgattatt gactagttat   15780
taatagtaat caattacggg gtcattagtt catagcccat atatggagtt ccgcgttaca   15840
taacttacgg taaatggccc gcctggctga ccgcccaacg accccgccc  attgacgtca   15900
ataatgacgt atgttcccat agtaacgcca atagggactt tccattgacg tcaatgggtg   15960
gagtatttac ggtaaactgc ccacttggca gtacatcaag tgtatcatat gccaagtccg   16020
cccccctattg acgtcaatga cggtaaatgg cccgcctggc attatgccca gtacatgacc   16080
ttacgggact ttcctacttg gcagtacatc tacgtattag tcatcgctat taccatggtg   16140
atgcggtttt ggcagtacat caatgggcgt ggatagcggt ttgactcacg gggatttcca   16200
agtctccacc ccattgacgt caatgggagt ttgttttggc accaaaatca acgggacttt   16260
ccaaaatgtc gtaacaactc cgccccattg acgcaaatgg cggtaggcg  tgtacggtgg   16320
gaggtctata taagcagagc tcgtttagtg aaccgtcaga tcgcctggag acgccatcca   16380
cgctgttttg acctccatag aagacaccgg gaccgatcca gcctcccctc gaagctcgac   16440
tctaggggct cgagatctgc gatctaagta agcttgcatg cctgcaggtc ggccgccacg   16500
accggtgccg ccaccatccc ctgacccacg cccctgaccc ctcacaagga gacgaccttc   16560
catgaccgag tacaagccca cggtgcgcct cgccacccgc gacgacgtcc cccgggccgt   16620
acgcaccctc gccgccgcgt tcgccgacta ccccgccacg cgccacaccg tcgacccgga   16680
ccgccacatc gagcgggtca ccgagctgca agaactcttc ctcacgcgcg tcgggctcga   16740
catcggcaag gtgtgggtcg cggacgacgg cgccgcggtg gcggtctgga ccacgccgga   16800
gagcgtcgaa gcggggcgg  tgttcgccga gatcggcccg cgcatggccg agttgagcgg   16860
ttcccggctg gccgcgcagc aacagatgga aggcctcctg gcgccgcacc ggcccaagga   16920
gcccgcgtgg ttcctggcca ccgtcggcgt ctcgcccgac caccagggca agggtctggg   16980
cagcgccgtc gtgctccccg gagtggaggc ggccgagcgc gccggggtgc cgccttcct   17040
ggagacctcc gcgccccgca acctcccctt ctacgagcgg ctcggcttca ccgtcaccgc   17100
cgacgtcgag gtgcccgaag accgcgcac  ctggtgcatg acccgcaagc ccggtgcctg   17160
acgcccgccc cacgacccgc agcgcccgac cgaaaggagc gcacgacccc atggctccga   17220
ccgaagccga cccgggcggc cccgccgacc ccgcacccgc ccccgaggcc caccgactct   17280
agagtcgggg cggccggccg cttcgagcag acatgataag atacattgat gagtttggac   17340
aaaccacaac tagaatgcag tgaaaaaaat gctttatttg tgaaatttgt gatgctattg   17400
ctttatttgt aaccattata agctgcaata aacaagttaa caacaacaat tgcattcatt   17460
ttatgtttca ggttcagggg gaggtgtggg aggttttta  aagcaagtaa aacctctaca   17520
aatgtggtaa aatcgataag gatcaattcg gcttcaggta ccgtcgacga tgtaggtcac   17580
ggtctcgaag ccgcggtgcg ggtgccaggg cgtgcccttg ggctccccgg gcgcgtactc   17640
cacctcaccc atctggtcca tcatgatgaa cgggtcgagg tggcggtagt tgatcccggc   17700
```

Page 26

```
                              Integrase.ST25
gaacgcgcgg cgcaccggga agccctcgcc ctcgaaaccg ctgggcgcgg tggtcacggt  17760

gagcacggga cgtgcgacgg cgtcggcggg tgcggatacg cggggcagcg tcagcgggtt  17820

ctcgacggtc acggcgggca tgtcgacagc cgaattgatc cgtcgaccga tgcccttgag  17880

agccttcaac ccagtcagct ccttccggtg ggcgcggggc atgactatcg tcgccgcact  17940

tatgactgtc ttctttatca tgcaactcgt aggacaggtg ccggcagcgc tcttccgctt  18000

cctcgctcac tgactcgctg cgctcggtcg ttcggctgcg gcgagcggta tcagctcact  18060

caaaggcggt aatacggtta tccacagaat caggggataa cgcaggaaag aacatg      18116


<210>  8
<211>  17402
<212>  DNA
<213>  Plasmid pOMIFN-Ins-CMV-pur-attB

<400>  8
ggccgccacc gcggtggagc tccaattcgc cctatagtga gtcgtattac aattcactgg     60

ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt tacccaactt aatcgccttg    120

cagcacatcc cccttttcgcc agctggcgta atagcgaaga ggcccgcacc gatcgccctt    180

cccaacagtt gcgcagcctg aatggcgaat gggacgcgcc ctgtagcggc gcattaagcg    240

cggcgggtgt ggtggttacg cgcagcgtga ccgctacact tgccagcgcc ctagcgcccg    300

ctcctttcgc tttcttccct tcctttctcg ccacgttcgc cggctttccc cgtcaagctc    360

taaatcgggg gctcccttta gggttccgat ttagtgcttt acggcacctc gaccccaaaa    420

aacttgatta gggtgatggt tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc    480

ctttgacgtt ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac    540

tcaaccctat ctcggtctat tcttttgatt tataagggat tttgccgatt tcggcctatt    600

ggttaaaaaa tgagctgatt taacaaaaat ttaacgcgaa ttttaacaaa atattaacgc    660

ttacaattta ggtggcactt ttcggggaaa tgtgcgcgga acccctattt gtttattttt    720

ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata    780

atattgaaaa aggaagagta tgagtattca acatttccgt gtcgccctta ttcccttttt    840

tgcggcattt tgccttcctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc    900

tgaagatcag ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaagat    960

ccttgagagt tttcgccccg aagaacgttt tccaatgatg agcactttta aagttctgct   1020

atgtggcgcg gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca   1080

ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg   1140

catgacagta agagaattat gcagtgctgc cataaccatg agtgataaca ctgcggccaa   1200

cttacttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg   1260

ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga   1320

cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg   1380

cgaactactt actctagctt cccggcaaca attaatagac tggatggagg cggataaagt   1440
```
                                  Page 27

Integrase.ST25

```
tgcaggacca cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg   1500

agccggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc   1560

ccgtatcgta gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca   1620

gatcgctgag ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc   1680

atatatactt tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat   1740

cctttttgat aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc   1800

agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg   1860

ctgcttgcaa acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct   1920

accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct   1980

tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct   2040

cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg   2100

gttggactca agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc   2160

gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga   2220

gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg   2280

cagggtcgga acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta   2340

tagtcctgtc gggtttcgcc acctctgact tgagcgtcga ttttgtgat gctcgtcagg   2400

ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc ggcctttttg   2460

ctggcctttt gctcacatgt tctttcctgc gttatcccct gattctgtgg ataaccgtat   2520

taccgccttt gagtgagctg ataccgctcg ccgcagccga acgaccgagc gcagcgagtc   2580

agtgagcgag gaagcggaag agcgcccaat acgcaaaccg cctctccccg cgcgttggcc   2640

gattcattaa tgcagctggc acgacaggtt cccgactgg aaagcgggca gtgagcgcaa   2700

cgcaattaat gtgagttagc tcactcatta ggcaccccag ctttacact ttatgcttcc   2760

ggctcgtatg ttgtgtggaa ttgtgagcgg ataacaattt cacacaggaa acagctatga   2820

ccatgattac gccaagctcg aaattaaccc tcactaaagg gaacaaaagc tgggtaccgg   2880

gccccccctc gactagaggg acagcccccc cccaaagccc ccagggatgt aattacgtcc   2940

ctcccccgct aggggcagc agcgagccgc ccggggctcc gctccggtcc ggcgctcccc   3000

ccgcatcccc gagccggcag cgtgcgggga cagcccgggc acggggaagg tggcacggga   3060

tcgctttcct ctgaacgctt ctcgctgctc tttgagcctg cagacacctg ggggatacg   3120

gggaaaaagc tttaggctga aagagagatt tagaatgaca gaatcataga acggcctggg   3180

ttgcaaagga gcacagtgct catccagatc caaccccctg ctatgtgcag ggtcatcaac   3240

cagcagccca ggctgcccag agccacatcc agcctggcct tgaatgcctg cagggatggg   3300

gcatccacag cctccttggg caacctgttc agtgcgtcac caccctctgg gggaaaaact   3360

gcctcctcat atccaaccca aacctcccct gtctcagtgt aaagccattc ccccttgtcc   3420

tatcaagggg gagtttgctg tgacattgtt ggtctggggt gacacatgtt tgccaattca   3480
```
Page 28

Integrase.ST25

```
gtgcatcacg gagaggcaga tcttggggat aaggaagtgc aggacagcat ggacgtggga    3540

catgcaggtg ttgagggctc tgggacactc tccaagtcac agcgttcaga acagccttaa    3600

ggataagaag ataggataga aggacaaaga gcaagttaaa acccagcatg gagaggagca    3660

caaaaaggcc acagacactg ctggtccctg tgtctgagcc tgcatgtttg atggtgtctg    3720

gatgcaagca gaaggggtgg aagagcttgc ctggagagat acagctgggt cagtaggact    3780

gggacaggca gctggagaat tgccatgtag atgttcatac aatcgtcaaa tcatgaaggc    3840

tggaaaagcc ctccaagatc cccaagacca accccaaccc acccaccgtg cccactggcc    3900

atgtccctca gtgccacatc cccacagttc ttcatcacct ccagggacgg tgaccccccc    3960

acctccgtgg gcagctgtgc cactgcagca ccgctctttg agaaggtaa atcttgctaa    4020

atccagcccg accctcccct ggcacaacgt aaggccatta tctctcatcc aactccagga    4080

cggagtcagt gaggatgggg ctctagtcga ggtcgacggt atcgataagc ttgattaggc    4140

agagcaatag gactctcaac ctcgtgagta tggcagcatg ttaactctgc actggagtcc    4200

agcgtgggaa acaatctgcc ttgcacatga gtcttcgtgg gccaatattc cccaacggtt    4260

ttccttcagc ttgtcttgtc tcctaagctc tcaaaacacc tttttggtga ataaactcac    4320

ttggcaacgt ttatctgtct taccttagtg tcacgtttca tccctattcc cctttctcct    4380

cctccgtgtg gtacacagtg gtgcacactg gttcttctgt tgatgttctg ctctgacagc    4440

caatgtgggt aaagttcttc ctgccacgtg tctgtgttgt tttcacttca aaaagggccc    4500

tgggctcccc ttggagctct caggcatttc cttaatcatc acagtcacgc tggcaggatt    4560

agtccctcct aaaccttaga atgacctgaa cgtgtgctcc ctctttgtag tcagtgcagg    4620

gagacgtttg cctcaagatc agggtccatc tcacccacag ggccattccc aagatgaggt    4680

ggatggttta ctctcacaaa aagttttctt atgtttggct agaaaggaga actcactgcc    4740

tacctgtgaa ttcccctagt cctggttctg ctgccactgc tgcctgtgca gcctgtccca    4800

tggaggggggc agcaactgct gtcacaaagg tgatcccacc ctgtctccac tgaaatgacc    4860

tcagtgccac gtgttgtata gggtataaag tacgggaggg ggatgcccgg ctcccttcag    4920

ggttgcagag cagaagtgtc tgtgtataga gtgtgtctta atctattaat gtaacagaac    4980

aacttcagtc ctagtgtttt gtgggctgga attgcccatg tggtagggac aggcctgcta    5040

aatcactgca atcgcctatg ttctgaaggt atttgggaaa gaaagggatt tgggggattg    5100

cctgtgattg gctttaattg aatggcaaat cacaggaaag cagttctgct caacagttgg    5160

ttgtttcagc caattcttgc agccaaagag ccgggtgccc agcgatataa tagttgtcac    5220

ttgtgtctgt atggatgaca gggaggtagg gtgacctgag gaccaccctc cagcttctgc    5280

tagcgtaggt acagtcacca cctccagctc cacacgagtc ccatcgtggt ttaccaaaga    5340

aacacaatta tttggaccag tttggaaagt cacccgctga attgtgaggc tagattaata    5400

gagctgaaga gcaaatgttc ccaacttgga gatactagtt ggtattagta tcagaggaac    5460

agggccatag cacctccatg ctattagatt ccggctggca tgtacttttc aagatgattt    5520
```

Integrase.ST25

```
gtaactaaca atggcttatt gtgcttgtct taagtctgtg tcctaatgta aatgttcctt     5580

tggtttatat aaccttcttg ccatttgctc ttcaggtgtt cttgcagaac actggctgct     5640

ttaatctagt ttaactgttg cttgattatt cttagggata agatctgaat aaactttttg     5700

tggctttggc agactttagc ttgggcttag ctcccacatt agcttttgct gccttttctg     5760

tgaagctatc aagatcctac tcaatgacat tagctgggtg caggtgtacc aaatcctgct     5820

ctgtggaaca cattgtctga tgataccgaa ggcaaacgtg aactcaaaga ggcacagagt     5880

taagaagaag tctgtgcaat tcagaggaaa agccaaagtg gccattagac acactttcca     5940

tgcagcattt gccagtaggt ttcatataaa actacaaaat ggaataaacc actacaaatg     6000

ggaaaagcct gatactagaa tttaaatatt cacccaggct caaggggtgt ttcatggagt     6060

aatatcactc tataaaagta gggcagccaa ttattcacag acaaagcttt ttttttctg      6120

tgctgcagtg ctgttttcg gctgatccag ggttacttat tgtgggtctg agagctgaat      6180

gatttctcct tgtgtcatgt tggtgaagga gatatggcca gggggagatg agcatgttca      6240

agaggaaacg ttgcattttg gtggcttggg agaaaggtag aacgatatca ggtccatagt      6300

gtcactaaga gatctgaagg atggttttac agaacagttg acttggctgg gtgcaggctt      6360

ggctgtaaat ggatggaagg atggacagat gggtggacag agatttctgt gcaggagatc      6420

atctcctgag ctcggtgctt gacagactgc agatccatcc cataaccttc tccagcatga      6480

gagcgcgggg agctttggta ctgttcagtc tgctgcttgt tgcttcctgg gtgcacagtg      6540

gtgattttct tactcacaca gggcaaaaac ctgagcagct tcaaagtgaa caggttgctc      6600

tcataggcca ttcagttgtc aagatgaggt ttttggtttc ttgttttgta aggtgggaag      6660

aagcactgaa ggatcagttg cgagggcagg ggtttagcac tgttcagaga agtcttattt      6720

taactcctct catgaacaaa aagagatgca ggtgcagatt ctggcaagca tgcagtgaag      6780

gagaaagccc tgaatttctg atatatgtgc aatgttgggc acctaacatt ccccgctgaa      6840

gcacagcagc tccagctcca tgcagtactc acagctggtg cagccctcgg ctccagggtc      6900

tgagcagtgc tgggactcac gaggttccat gtctttcaca ctgataatgg tccaatttct      6960

ggaatgggtg cccatccttg gaggtcccca aggccaggct ggctgcgtct ccgagcagcc      7020

cgatctggtg gtgagtagcc agcccatggc aggagttaga gcctgatggt cttttaaggtc     7080

ccttccaacc taagccatcc tacgattcta ggaatcatga cttgtgagtg tgtattgcag      7140

aggcaatatt ttaaagttat aaatgttttc tcccttcct tgtttgtcaa agttatcttg       7200

atcgccttat caatgctttt ggagtctcca gtcatttttc ttacamcaaa aagaggagga     7260

agaatgaaga gaatcattta atttcttgat tgaatagtag gattcagaaa gctgtacgta     7320

atgccgtctc tttgtatcga gctgtaaggt ttctcatcat ttatcagcgt ggtacatatc     7380

agcactttc catctgatgt ggaaaaaaaa atccttatca tctacagtct ctgtacctaa      7440

acatcgctca gactctttac caaaaaagct ataggtttta aaactacatc tgctgataat     7500

ttgccttgtt ttagctcttc ttccatatgc tgcgtttgtg agaggtgcgt ggatgggcct     7560
```

Page 30

Integrase.ST25

aaactctcag	ctgctgagct	tgatgggtgc	ttaagaatga	agcactcact	gctgaaactg	7620
ttttcatttc	acaggaatgt	tttagtggca	ttgttttat	aactacatat	tcctcagata	7680
aatgaaatcc	agaaataatt	atgcaaactc	actgcatccg	ttgcacaggt	ctttatctgc	7740
tagcaaagga	aataatttgg	ggatggcaaa	aacattcctt	cagacatcta	tatttaaagg	7800
aatataatcc	tggtacccac	ccacttcatc	cctcattatg	ttcacactca	gagatactca	7860
ttctcttgtt	gttatcattt	gatagcgttt	tctttggttc	tttgccacgc	tctgggctat	7920
ggctgcacgc	tctgcactga	tcagcaagta	gatgcgaggg	aagcagcagt	gagaggggct	7980
gccctcagct	ggcacccagc	cgctcagcct	aggaggggac	cttgcctttc	caccagctga	8040
ggtgcagccc	tacaagctta	cacgtgctgc	gagcaggtga	gcaaagggag	tcttcatggt	8100
gtgtttcttg	ctgcccggaa	gcaaaacttt	actttcattc	attccccttg	aagaatgagg	8160
aatgtttgga	aacggactgc	tttacgttca	atttctctct	tccctttaag	gctcagccag	8220
gggccattgc	tgaggacggc	atcggggccc	cctggaccaa	atctgtggca	cagatggttt	8280
cacttacatc	agtggatgtg	ggatctgcgc	ctgtaatgtg	tccttctgaa	ggaaggaacg	8340
tgccttccaa	gtgccagccc	cacagccccc	agcccctccc	tgtgctgctc	caattcatct	8400
cctcttcctc	cttctccctt	tgctgtttgt	gctcgggtag	aaatcatgaa	gatttagaag	8460
agaaaacaaa	ataactggag	tggaaaccca	ggtgatgcag	ttcattcagc	tgtcataggt	8520
ttgtcgttgc	tataggtctg	tatcagagat	gctarcacca	ctttgctgtc	ggtgcttaac	8580
tcgggtgaac	tctccttcac	tcgcatcatt	tgcgggcctt	atttacatcc	ccagcatcca	8640
tcaccctctg	ggaaatgggg	cgcactggat	ctctaatgga	agactttccc	tctttcagag	8700
cctgtgggat	gtgcagtgac	aagaaacgtg	gaggggctga	gcagcagcac	tgcccccagg	8760
gagcaggagc	ggatgccatc	ggtggcagca	tcccaaatga	tgtcagcgga	tgctgagcag	8820
gcagcggacg	aacggacaga	agcgatgcgt	acaccttctg	ttgacatggt	atttggcagc	8880
gatttaacac	tcgcttccta	gtcctgctat	tctccacagg	ctgcattcaa	atgaacgaag	8940
ggaagggagg	caaaaagatg	caaatccga	dacaagcagc	agaaatattt	cttcgctacg	9000
gaagcgtgcg	caaacaacct	tctccaacag	caccagaaga	gcacagcgta	accttttca	9060
agaccagaaa	aggaaattca	caaagcctct	gtggatacca	gcgcgttcag	ctctcctgat	9120
agcagatttc	ttgtcaggtt	gcgaatgggg	tatggtgcca	ggaggtgcag	ggaccatatg	9180
atcatataca	gcacagcagt	cattgtgcat	gtattaatat	atattgagta	gcagtgttac	9240
tttgccaaag	caatagttca	gagatgagtc	ctgctgcata	cctctatctt	aaaactaact	9300
tataaatagt	aaaaccttct	cagttcagcc	acgtgctcct	ctctgtcagc	accaatggtg	9360
cttcgcctgc	acccagctgc	aaggaatcag	cccgtgatct	cattaacact	cagctctgca	9420
ggataaatta	gattgttcca	ctctcttttg	ttgttaatta	cgacggaaca	attgttcagt	9480
gctgatggtc	ctaattgtca	gctacagaaa	acgtctccat	gcagttcctt	ctgcgccagc	9540
aaactgtcca	ggctatagca	ccgtgatgca	tgctacctct	cactccatcc	ttcttctctt	9600

Page 31

Integrase.ST25

```
tcccaccagg gagagctgtg tgttttcact ctcagccact ctgaacaata ccaaactgct    9660

acgcactgcc tccctcggaa agagaatccc cttgttgctt ttttatttac aggatccttc    9720

ttaaaaagca gaccatcatt cactgcaaac ccagagcttc atgcctctcc ttccacaacc    9780

gaaaacagcc ggcttcattt gtcttttta aatgctgttt tccaggtgaa ttttggccag    9840

cgtgttggct gagatccagg agcacgtgtc agctttctgc tctcattgct cctgttctgc    9900

attgcctctt tctggggttt ccaagagggg gggagacttt gcgcggggat gagataatgc    9960

cccttttctt agggtggctg ctgggcagca gagtggctct gggtcactgt ggcaccaatg   10020

ggaggcacca gtgggggtgt gttttgtgca gggggggaagc attcacagaa tggggctgat   10080

cctgaagctt gcagtccaag gctttgtctg tgtacccagt gaaatccttc ctctgttaca   10140

taaagcccag ataggactca gaaatgtagt cattccagcc cccctcttcc tcagatctgg   10200

agcagcactt gtttgcagcc agtcctcccc aaaatgcaca gacctcgccg agtggaggga   10260

gatgtaaaca gcgaaggtta attacctcct tgtcaaaac actttgtggt ccatagatgt   10320

ttctgtcaat cttacaaaac agaaccgaga ggcagcgagc actgaagagc gtgttcccat   10380

gctgagttaa tgagacttgg cagctcgctg tgcagagatg atccctgtgc ttcatgggag   10440

gctgtaacct gtctccccat cgccttcaca ccgcagtgct gtcctggaca cctcaccctc   10500

cataagctgt aggatgcagc tgcccaggga tcaagagact tttcctaagg ctcttaggac   10560

tcatctttgc cgctcagtag cgtgcagcaa ttactcatcc caactatact gaatgggttt   10620

ctgccagctc tgcttgtttg tcaataagca tttcttcatt ttgcctctaa gtttctctca   10680

gcagcaccgc tctgggtgac ctgagtggcc acctggaacc cgaggggcac agccaccacc   10740

tccctgttgc tgctgctcca gggactcatg tgctgctgga tggggggaag catgaagttc   10800

ctcacccaga cacctgggtt gcaatggctg cagcgtgctc ttcttggtat gcagattgtt   10860

tccagccatt acttgtagaa atgtgctgtg gaagcccttt gtatctcttt ctgtggccct   10920

tcagcaaaag ctgtgggaaa gctctgaggc tgctttcttg ggtcgtggag gaattgtatg   10980

ttccttcttt aacaaaaatt atccttagga gagagcactg tgcaagcatt gtgcacataa   11040

aacaattcag gttgaaaggg ctctctggag gtttccagcc tgactactgc tcgaagcaag   11100

gccaggttca aagatggctc aggatgctgt gtgccttcct gattatctgt gccaccaatg   11160

gaggagattc acagccactc tgcttcccgt gccactcatg gagaggaata ttcccttata   11220

ttcagataga atgttatcct ttagctcagc cttccctata accccatgag ggagctgcag   11280

atccccatac tctccccttc tctggggtga aggccgtgtc ccccagcccc ccttcccacc   11340

ctgtgcccta agcagcccgc tggcctctgc tggatgtgtg cctatatgtc aatgcctgtc   11400

cttgcagtcc agcctgggac atttaattca tcaccagggt aatgtggaac tgtgtcatct   11460

tcccctgcag ggtacaaagt tctgcacggg gtcctttcgg ttcaggaaaa ccttcactgg   11520

tgctacctga atcaagctct atttaataag ttcataagca catggatgtg ttttcctaga   11580

gatacgtttt aatggtatca gtgattttta tttgctttgt tgcttacttc aaacagtgcc   11640
```

Page 32

Integrase.ST25

tttgggcagg	aggtgaggga	cgggtctgcc	gttggctctg	cagtgatttc	tccaggcgtg	11700
tggctcaggt	cagatagtgg	tcactctgtg	gccagaagaa	ggacaaagat	ggaaattgca	11760
gattgagtca	cgttaagcag	gcatcttgga	gtgatttgag	gcagtttcat	gaaagagcta	11820
cgaccactta	ttgttgtttt	ccccttttac	aacagaagtt	ttcatcaaaa	taacgtggca	11880
aagcccagga	atgtttggga	aaagtgtagt	taaatgtttt	gtaattcatt	tgtcggagtg	11940
ctaccagcta	agaaaaaagt	cctacctttg	gtatggtagt	cctgcagaga	atacaacatc	12000
aatattagtt	tggaaaaaaa	caccaccacc	accagaaact	gtaatggaaa	atgtaaacca	12060
agaaattcct	tgggtaagag	agaaaggatg	tcgtatactg	gccaagtcct	gcccagctgt	12120
cagcctgctg	accctctgca	gttcaggacc	atgaaacgtg	gcactgtaag	acgtgtcccc	12180
tgcctttgct	tgcccacaga	tctctgccct	gtgctgact	cctgcacaca	agagcatttc	12240
cctgtagcca	aacagcgatt	agccataagc	tgcacctgac	tttgaggatt	aagagtttgc	12300
aattaagtgg	attgcagcag	gagatcagtg	gcagggttgc	agatgaaatc	cttttctagg	12360
ggtagctaag	ggctgagcaa	cctgtcctac	agcacaagcc	aaaccagcca	agggttttcc	12420
tgtgctgttc	acagaggcag	ggccagctgg	agctggagga	ggttgtgctg	ggacccttct	12480
ccctgtgctg	agaatggagt	gatttctggg	tgctgttcct	gtggcttgca	ctgagcagct	12540
caagggagat	cggtgctcct	catgcagtgc	caaaactcgt	gtttgatgca	gaaagatgga	12600
tgtgcacctc	cctcctgcta	atgcagccgt	gagcttatga	aggcaatgag	ccctcagtgc	12660
agcaggagct	gtagtgcact	cctgtaggtg	ctagggaaaa	tctctggttc	cagggatgc	12720
attcataagg	gcaatatatc	ttgaggctgc	gccaaatctt	tctgaaatat	tcatgcgtgt	12780
tcccttaatt	tatagaaaca	aacacagcag	aataattatt	ccaatgcctc	ccctcgaagg	12840
aaacccatat	ttccatgtag	aaatgtaacc	tatatacaca	cagccatgct	gcatccttca	12900
gaacgtgcca	gtgctcatct	cccatggcaa	aatactacag	gtattctcac	tatgttggac	12960
ctgtgaaagg	aaccatggta	agaaacttcg	gttaaaggta	tggctgcaaa	actactcata	13020
ccaaaacagc	agagctccag	acctcctctt	aggaaagagc	cacttggaga	gggatggtgt	13080
gaaggctgga	ggtgagagac	agagcctgtc	ccagttttcc	tgtctctatt	ttctgaaacg	13140
tttgcaggag	gaaaggacaa	ctgtactttc	aggcatagct	ggtgccctca	cgtaaataag	13200
ttccccgaac	ttctgtgtca	tttgttctta	agatgctttg	gcagaacact	ttgagtcaat	13260
tcgcttaact	gtgactaggt	ctgtaaataa	gtgctccctg	ctgataaggt	tcaagtgaca	13320
ttttttagtgg	tatttgacag	catttacctt	gctttcaagt	cttctaccaa	gctcttctat	13380
acttaagcag	tgaaaccgcc	aagaaaccct	tccttttatc	aagctagtgc	taaataccat	13440
taacttcata	ggttagatac	ggtgctgcca	gcttcacctg	gcagtggttg	gtcagttctg	13500
ctggtgacaa	agcctccctg	gcctgtgctt	ttacctagag	gtgaatatcc	aagaatgcag	13560
aactgcatgg	aaagcagagc	tgcaggcacg	atggtgctga	gccttagctg	cttcctgctg	13620
ggagatgtgg	atgcagagac	gaatgaagga	cctgtccctt	actcccctca	gcattctgtg	13680

Integrase.ST25

```
ctatttaggg ttctaccaga gtccttaaga ggtttttttt ttttttggtc caaaagtctg   13740

tttgtttggt tttgaccact gagagcatgt gacacttgtc tcaagctatt aaccaagtgt   13800

ccagccaaaa tcaattgcct gggagacgca gaccattacc tggaggtcag gacctcaata   13860

aatattacca gcctcattgt gccgctgaca gattcagctg gctgctccgt gttccagtcc   13920

aacagttcgg acgccacgtt tgtatatatt tgcaggcagc ctcggggggga ccatctcagg   13980

agcagagcac cggcagccgc ctgcagagcc gggcagtacc tcaccatggc tttgaccttt   14040

gccttactgg tggctctcct ggtgctgagc tgcaagagca gctgctctgt gggctgcgat   14100

ctgcctcaga cccacagcct gggcagcagg aggaccctga tgctgctggc tcagatgagg   14160

agaatcagcc tgtttagctg cctgaaggat aggcacgatt ttggctttcc tcaagaggag   14220

tttggcaacc agtttcagaa ggctgagacc atccctgtgc tgcacgagat gatccagcag   14280

atctttaacc tgtttagcac caaggatagc agcgctgctt gggatgagac cctgctggat   14340

aagtttttaca ccgagctgta ccagcagctg aacgatctgg aggcttgcgt gatccagggc   14400

gtgggcgtga ccgagacccc tctgatgaag gaggatagca tcctggctgt gaggaagtac   14460

tttcagagga tcaccctgta cctgaaggag aagaagtaca gccctgcgc ttgggaagtc   14520

gtgagggctg agatcatgag gagctttagc ctgagcacca acctgcaaga gagcttgagg   14580

tctaaggagt aaaaagtcta gagtcggggc ggccggccgc ttcgagcaga catgataaga   14640

tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaaatg ctttatttgt   14700

gaaatttgtg atgctattgc tttatttgta accattataa gctgcaataa acaagttaac   14760

aacaacaatt gcattcattt tatgtttcag gttcagggggg aggtgtggga ggttttttaa   14820

agcaagtaaa acctctacaa atgtggtaaa atcgataccg tcgacctcga ctagagcggc   14880

cactaacata cgctctccat caaaacaaaa cgaaacaaaa caaactagca aaataggctg   14940

tccccagtgc aagtgcaggt gccagaacat ttctctatcg ataggtaccg agctcttacg   15000

cgtgctagcc ctcgagcagg atctatacat tgaatcaata ttggcaatta gccatattag   15060

tcattggtta tatagcataa atcaatattg gctattggcc attgcatacg ttgtatctat   15120

atcataatat gtacatttat attggctcat gtccaatatg accgccatgt tgacattgat   15180

tattgactag ttattaatag taatcaatta cggggtcatt agttcatagc ccatatatgg   15240

agttccgcgt tacataactt acggtaaatg cccgcctgg ctgaccgccc aacgaccccc   15300

gcccattgac gtcaataatg acgtatgttc ccatagtaac gccaataggg actttccatt   15360

gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt ggcagtacat caagtgtatc   15420

atatgccaag tccgcccct attgacgtca atgacggtaa atggcccgcc tggcattatg   15480

cccagtacat gaccttacgg gactttccta cttggcagta catctacgta ttagtcatcg   15540

ctattaccat ggtgatgcgg ttttggcagt acatcaatgg gcgtggatag cggtttgact   15600

cacgggatt tccaagtctc cacccccattg acgtcaatgg gagtttgttt tggcaccaaa   15660

atcaacggga ctttccaaaa tgtcgtaaca actccgcccc attgacgcaa atgggcggta   15720
```
Page 34

Integrase.ST25

```
ggcgtgtacg gtgggaggtc tatataagca gagctcgttt agtgaaccgt cagatcgcct    15780

ggagacgcca tccacgctgt tttgacctcc atagaagaca ccgggaccga tccagcctcc    15840

cctcgaagct cgactctagg ggctcgagat ctgcgatcta agtaagcttg catgcctgca    15900

ggtcggccgc cacgaccggt gccgccacca tcccctgacc cacgcccctg acccctcaca    15960

aggagacgac cttccatgac cgagtacaag cccacggtgc gcctcgccac ccgcgacgac    16020

gtcccccggg ccgtacgcac cctcgccgcc gcgttcgccg actacccgc cacgcgccac      16080

accgtcgacc cggaccgcca catcgagcgg gtcaccgagc tgcaagaact cttcctcacg    16140

cgcgtcgggc tcgacatcgg caaggtgtgg gtcgcggacg acggcgccgc ggtggcggtc    16200

tggaccacgc cggagagcgt cgaagcgggg gcggtgttcg ccgagatcgg cccgcgcatg    16260

gccgagttga gcggttcccg gctggccgcg cagcaacaga tggaaggcct cctggcgccg    16320

caccggccca aggagcccgc gtggttcctg gccaccgtcg gcgtctcgcc cgaccaccag    16380

ggcaagggtc tgggcagcgc cgtcgtgctc cccggagtgg aggcggccga gcgcgccggg    16440

gtgcccgcct tcctggagac ctccgcgccc cgcaacctcc ccttctacga gcggctcggc    16500

ttcaccgtca ccgccgacgt cgaggtgccc gaaggaccgc gcacctggtg catgacccgc    16560

aagcccggtg cctgacgccc gccccacgac ccgcagcgcc cgaccgaaag gagcgcacga    16620

ccccatggct ccgaccgaag ccgacccggg cggccccgcc gaccccgcac cgcccccga     16680

ggcccaccga ctctagagtc ggggcggccg gccgcttcga gcagacatga taagatacat    16740

tgatgagttt ggacaaacca caactagaat gcagtgaaaa aaatgcttta tttgtgaaat    16800

ttgtgatgct attgctttat ttgtaaccat tataagctgc aataaacaag ttaacaacaa    16860

caattgcatt cattttatgt ttcaggttca gggggaggtg tgggaggttt tttaaagcaa    16920

gtaaaacctc tacaaatgtg gtaaaatcga taaggatcaa ttcggcttca ggtaccgtcg    16980

acgatgtagg tcacggtctc gaagccgcgg tgcgggtgcc agggcgtgcc cttgggctcc    17040

ccgggcgcgt actccacctc acccatctgg tccatcatga tgaacgggtc gaggtggcgg    17100

tagttgatcc cggcgaacgc gcggcgcacc gggaagccct cgccctcgaa accgctgggc    17160

gcggtggtca cggtgagcac gggacgtgcg acggcgtcgg cgggtgcgga tacgcggggc    17220

agcgtcagcg ggttctcgac ggtcacggcg ggcatgtcga cagccgaatt gatccgtcga    17280

ccgatgccct tgagagcctt caacccagtc agctccttcc ggtgggcgcg ggcatgact      17340

atcgtcgccg cacttatgac tgtcttcttt atcatgcaac tcgtaggaca ggtgccggca     17400

gc                                                                   17402
```

<210> 9
<211> 5172
<212> DNA
<213> Plasmid pRSV-Int

<400> 9
ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg cgctcttcc   60

Integrase.ST25

```
gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct    120

cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg    180

tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc    240

cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga    300

aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct    360

cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg    420

gcgctttctc aatgctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag    480

ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat    540

cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac    600

aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac    660

tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc    720

ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt    780

tttgtttgca agcagcagat tacgcgcaga aaaaaggat ctcaagaaga tcctttgatc    840

ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg    900

agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca    960

atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca   1020

cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag   1080

ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac   1140

ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc   1200

agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct   1260

agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc   1320

gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg   1380

cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc   1440

gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat   1500

tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag   1560

tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat   1620

aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg   1680

cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca   1740

cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga   1800

aggcaaaatg ccgcaaaaaa gggaataagg cgacacgga atgttgaat actcatactc   1860

ttcctttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata   1920

tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg   1980

ccacctgacg tcgacggatc gggagatctc ccgatcccct atggtcgact ctcagtacaa   2040

tctgctctga tgccgcatag ttaagccagt atctgctccc tgcttgtgtg ttggaggtcg   2100
```

Integrase.ST25

```
ctgagtagtg cgcgagcaaa atttaagcta caacaaggca aggcttgacc gacaattgca  2160
tgaagaatct gcttagggtt aggcgttttg cgctgcttcg cgatgtacgg gccagatata  2220
cgcgtgctag gggtctagga tcgattctag gaattctcta gccgcggtct agggatcccg  2280
gcgcgtatgg tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagtatct  2340
gctccctgct tgtgtgttgg aggtcgctga gtagtgcgcg agcaaaattt aagctacaac  2400
aaggcaaggc ttgaccgaca attgcatgaa gaatctgctt agggttaggc gttttgcgct  2460
gcttcgcgat gtacgggcca gatatacgcg tatctgaggg gactagggtg tgtttaggcg  2520
aaaagcgggg cttcggttgt acgcggttag gagtcccctc aggatatagt agtttcgctt  2580
ttgcataggg aggggggaaat gtagtcttat gcaatacact tgtagtcttg caacatggta  2640
acgatgagtt agcaacatgc cttacaagga gagaaaaagc accgtgcatg ccgattggtg  2700
gaagtaaggt ggtacgatcg tgccttatta ggaaggcaac agacaggtct gacatggatt  2760
ggacgaacca ctgaattccg cattgcagag ataattgtat ttaagtgcct agctcgatac  2820
aataaacgcc atttgaccat tcaccacatt ggtgtgcacc tccaagcttg catgcctgca  2880
ggtaccggtc cggaattccc gggtcgacga gctcactagt cgtagggtcg ccgacatgac  2940
acaagggggtt gtgaccgggg tggacacgta cgcgggtgct tacgaccgtc agtcgcgcga  3000
gcgcgagaat tcgagcgcag caagcccagc gacacagcgt agcgccaacg aagacaaggc  3060
ggccgacctt cagcgcgaag tcgagcgcga cggggggccgg ttcaggttcg tcgggcattt  3120
cagcgaagcg ccgggcacgt cggcgttcgg gacggcggag cgcccggagt tcgaacgcat  3180
cctgaacgaa tgccgcgccg ggcggctcaa catgatcatt gtctatgacg tgtcgcgctt  3240
ctcgcgcctg aaggtcatgg acgcgattcc gattgtctcg gaattgctcg ccctgggcgt  3300
gacgattgtt tccactcagg aaggcgtctt ccggcaggga aacgtcatgg acctgattca  3360
cctgattatg cggctcgacg cgtcgcacaa agaatcttcg ctgaagtcgg cgaagattct  3420
cgacacgaag aaccttcagc gcgaattggg cgggtacgtc ggcgggaagg cgccttacgg  3480
cttcgagctt gtttcggaga cgaaggagat cacgcgcaac ggccgaatgg tcaatgtcgt  3540
catcaacaag cttgcgcact cgaccactcc ccttaccgga cccttcgagt tcgagcccga  3600
cgtaatccgg tggtggtggc gtgagatcaa gacgcacaaa caccttccct tcaagccggg  3660
cagtcaagcc gccattcacc cgggcagcat cacggggctt tgtaagcgca tggacgctga  3720
cgccgtgccg acccggggcg agacgattgg gaagaagacc gcttcaagcg cctgggaccc  3780
ggcaaccgtt atgcgaatcc ttcgggaccc gcgtattgcg ggcttcgccg ctgaggtgat  3840
ctacaagaag aagccggacg gcacgccgac cacgaagatt gagggttacc gcattcagcg  3900
cgacccgatc acgctccggc cggtcgagct tgattgcgga ccgatcatcg agcccgctga  3960
gtggtatgag cttcaggcgt ggttggacgg cagggggcgc ggcaagggggc tttcccgggg  4020
gcaagccatt ctgtccgcca tggacaagct gtactgcgag tgtggcgccg tcatgacttc  4080
gaagcgcggg gaagaatcga tcaaggactc ttaccgctgc cgtcgccgga aggtggtcga  4140
```

Page 37

Integrase.ST25

```
cccgtccgca cctgggcagc acgaaggcac gtgcaacgtc agcatggcgg cactcgacaa     4200
gttcgttgcg gaacgcatct tcaacaagat caggcacgcc gaaggcgacg aagagacgtt     4260
ggcgcttctg tgggaagccg cccgacgctt cggcaagctc actgaggcgc ctgagaagag     4320
cggcgaacgg gcgaaccttg ttgcggagcg cgccgacgcc ctgaacgccc ttgaagagct     4380
gtacgaagac cgcgcggcag gcgcgtacga cggacccgtt ggcaggaagc acttccggaa     4440
gcaacaggca gcgctgacgc tccggcagca aggggcggaa gagcggcttg ccgaacttga     4500
agccgccgaa gccccgaagc ttccccttga ccaatggttc cccgaagacg ccgacgctga     4560
cccgaccggc cctaagtcgt ggtgggggcg cgcgtcagta gacgacaagc gcgtgttcgt     4620
cgggctcttc gtagacaaga tcgttgtcac gaagtcgact acgggcaggg ggcagggaac     4680
gcccatcgag aagcgcgctt cgatcacgtg ggcgaagccg ccgaccgacg acgacgaaga     4740
cgacgcccag gacggcacgg aagacgtagc ggcgtagcga gacacccgga tccctcgagg     4800
ggccctattc tatagtgtca cctaaatgct agagctcgct gatcagcctc gactgtgcct     4860
tctagttgcc agccatctgt tgtttgcccc tcccccgtgc cttccttgac cctggaaggt     4920
gccactccca ctgtcctttc ctaataaaat gaggaaattg catcgcattg tctgagtagg     4980
tgtcattcta ttctggggg g tggggtgggg caggacagca aggggggagga ttgggaagac     5040
aatagcaggc atgctgggga tgcggtgggc tctatggctt ctgaggcgga aagaaccagg     5100
tgcccagtca tagccgaata gcctctccac ccaagcggcc ggagaacctg cgtgcaatcc     5160
actggggggcg cg                                                         5172
```

<210>   10
<211>   6233
<212>   DNA
<213>   Plasmid pCR-XL-TOPO-CMV-pur-attB

<400>   10
```
agcgcccaat acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc      60
acgacaggtt tcccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc     120
tcactcatta ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa     180
ttgtgagcgg ataacaattt cacacaggaa acagctatga ccatgattac gccaagctat     240
ttaggtgacg cgttagaata ctcaagctat gcatcaagct tggtaccgag ctcggatcca     300
ctagtaacgg ccgccagtgt gctggaattc gcccttggcc gcataaaat atctttattt     360
tcattcatc tgtgtgttgg tttttgtgt gaatcgatag tactaacata cgctctccat     420
caaaacaaaa cgaaacaaaa caaactagca aaataggctg tccccagtgc aagtgcaggt     480
gccagaacat ttctctatcg ataggtaccg agctcttacg cgtgctagcc ctcgagcagg     540
atctatacat tgaatcaata ttggcaatta gccatattag tcattggtta tatagcataa     600
atcaatattg gctattggcc attgcatacg ttgtatctat atcataatat gtacatttat     660
attggctcat gtccaatatg accgccatgt tgacattgat tattgactag ttattaatag     720
taatcaatta cggggtcatt agttcatagc ccatatatgg agttccgcgt tacataactt     780
```

Integrase.ST25

```
acggtaaatg gcccgcctgg ctgaccgccc aacgaccccc gcccattgac gtcaataatg    840

acgtatgttc ccatagtaac gccaataggg actttccatt gacgtcaatg ggtggagtat    900

ttacggtaaa ctgcccactt ggcagtacat caagtgtatc atatgccaag tccgccccct    960

attgacgtca atgacggtaa atggcccgcc tggcattatg cccagtacat gaccttacgg   1020

gactttccta cttggcagta catctacgta ttagtcatcg ctattaccat ggtgatgcgg   1080

ttttggcagt acatcaatgg gcgtggatag cggtttgact cacggggatt tccaagtctc   1140

caccccattg acgtcaatgg gagtttgttt tggcaccaaa atcaacggga ctttccaaaa   1200

tgtcgtaaca actccgcccc attgacgcaa atgggcggta ggcgtgtacg gtgggaggtc   1260

tatataagca gagctcgttt agtgaaccgt cagatcgcct ggagacgcca tccacgctgt   1320

tttgacctcc atagaagaca ccgggaccga tccagcctcc cctcgaagct cgactctagg   1380

ggctcgagat ctgcgatcta agtaagcttg catgcctgca ggtcggccgc cacgaccggt   1440

gccgccacca tcccctgacc cacgccctg acccctcaca aggagacgac cttccatgac   1500

cgagtacaag cccacggtgc gcctcgccac ccgcgacgac gtccccggg ccgtacgcac   1560

cctcgccgcc gcgttcgccg actaccccgc cacgcgccac accgtcgacc cggaccgcca   1620

catcgagcgg gtcaccgagc tgcaagaact cttcctcacg cgcgtcgggc tcgacatcgg   1680

caaggtgtgg gtcgcggacg acggcgccgc ggtggcggtc tggaccacgc cggagagcgt   1740

cgaagcgggg gcggtgttcg ccgagatcgg cccgcgcatg gccgagttga gcggttcccg   1800

gctggccgcg cagcaacaga tggaaggcct cctggcgccg caccggccca aggagcccgc   1860

gtggttcctg gccaccgtcg gcgtctcgcc cgaccaccag ggcaagggtc tgggcagcgc   1920

cgtcgtgctc cccggagtgg aggcggccga gcgcgccggg gtgcccgcct cctggagac   1980

ctccgcgccc cgcaacctcc ccttctacga gcggctcggc ttcaccgtca ccgccgacgt   2040

cgaggtgccc gaaggaccgc gcacctggtg catgacccgc aagcccggtg cctgacgccc   2100

gcccacgac ccgcagcgcc cgaccgaaag gagcgcacga ccccatggct ccgaccgaag   2160

ccgacccggg cggccccgcc gaccccgcac ccgcccccga ggcccaccga ctctagagtc   2220

ggggcggccg gccgcttcga gcagacatga taagatacat tgatgagttt ggacaaacca   2280

caactagaat gcagtgaaaa aaatgcttta tttgtgaaat ttgtgatgct attgctttat   2340

ttgtaaccat tataagctgc aataaacaag ttaacaacaa caattgcatt catttttatgt   2400

ttcaggttca ggggggaggtg tgggaggttt tttaaagcaa gtaaaacctc tacaaatgtg   2460

gtaaaatcga taaggatcaa ttcggcttca ggtaccgtcg acgatgtagg tcacggtctc   2520

gaagccgcgg tgcgggtgcc agggcgtgcc cttgggctcc ccgggcgcgt actccacctc   2580

acccatctgg tccatcatga tgaacgggtc gaggtggcgg tagttgatcc ggcgaacgc   2640

gcggcgcacc gggaagccct cgccctcgaa accgctgggc gcggtggtca cggtgagcac   2700

gggacgtgcg acggcgtcgg cgggtgcgga tacgcggggc agcgtcagcg ggttctcgac   2760

ggtcacggcg ggcatgtcga cagccgaatt gatccgtcga ccgatgccct tgagagcctt   2820
```

Integrase.ST25

```
caacccagtc agctccttcc ggtgggcgcg gggcatgact atcgtcgccg cacttatgac      2880

tgtcttcttt atcatgcaac tcgtaggaca ggtgccggca gcgctcttcc gcttcctcgc      2940

tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg      3000

cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg aagggcgaat      3060

tctgcagata tccatcacac tggcggccgc tcgagcatgc atctagaggg cccaattcgc      3120

cctatagtga gtcgtattac aattcactgg ccgtcgtttt acaacgtcgt gactgggaaa      3180

accctggcgt tacccaactt aatcgccttg cagcacatcc ccctttcgcc agctggcgta      3240

atagcgaaga ggcccgcacc gatcgccctt cccaacagtt gcgcagccta tacgtacggc      3300

agtttaaggt ttacacctat aaaagagaga ccgttatcg tctgtttgtg gatgtacaga      3360

gtgatattat tgacacgccg gggcgacgga tggtgatccc cctggccagt gcacgtctgc      3420

tgtcagataa agtctcccgt gaactttacc cggtggtgca tatcggggat gaaagctggc      3480

gcatgatgac caccgatatg gccagtgtgc cggtctccgt tatcggggaa gaagtggctg      3540

atctcagcca ccgcgaaaat gacatcaaaa acgccattaa cctgatgttc tggggaatat      3600

aaatgtcagg catgagatta tcaaaaagga tcttcaccta gatccttttc acgtagaaag      3660

ccagtccgca gaaacggtgc tgaccccgga tgaatgtcag ctactgggct atctggacaa      3720

gggaaaacgc aagcgcaaag agaaagcagg tagcttgcag tgggcttaca tggcgatagc      3780

tagactgggc ggttttatgg acagcaagcg aaccggaatt gccagctggg cgccctctg      3840

gtaaggttgg gaagccctgc aaagtaaact ggatggcttt ctcgccgcca aggatctgat      3900

ggcgcagggg atcaagctct gatcaagaga caggatgagg atcgtttcgc atgattgaac      3960

aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc ggctatgact      4020

gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca gcgcaggggc      4080

gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg caagacgagg      4140

cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcagctgtg ctcgacgttg      4200

tcactgaagc gggaagggac tggctgctat tgggcgaagt gccggggcag gatctcctgt      4260

catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc      4320

atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcgag      4380

cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa gagcatcagg      4440

ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgag catgcccgac ggcgaggatc      4500

tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt      4560

ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg      4620

ctacccgtga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctcgtgcttt      4680

acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagttct      4740

tctgaattat taacgcttac aatttcctga tgcggtattt tctccttacg catctgtgcg      4800

gtatttcaca ccgcatacag gtggcacttt tcggggaaat gtgcgcggaa cccctatttg      4860
```

Integrase.ST25

```
tttattttc  taaatacatt  caaatatgta  tccgctcatg  agacaataac  cctgataaat   4920

gcttcaataa  tagcacgtga  ggagggccac  catggccaag  ttgaccagtg  ccgttccggt   4980

gctcaccgcg  cgcgacgtcg  ccggagcggt  cgagttctgg  accgaccggc  tcgggttctc   5040

ccgggacttc  gtggaggacg  acttcgccgg  tgtggtccgg  acgacgtga  ccctgttcat   5100

cagcgcggtc  caggaccagg  tggtgccgga  caacaccctg  gcctgggtgt  gggtgcgcgg   5160

cctggacgag  ctgtacgccg  agtggtcgga  ggtcgtgtcc  acgaacttcc  gggacgcctc   5220

cgggccggcc  atgaccgaga  tcggcgagca  gccgtggggg  cgggagttcg  ccctgcgcga   5280

cccggccggc  aactgcgtgc  acttcgtggc  cgaggagcag  gactgacacg  tgctaaaact   5340

tcattttaa  tttaaaagga  tctaggtgaa  gatcctttt  gataatctca  tgaccaaaat   5400

cccttaacgt  gagttttcgt  tccactgagc  gtcagacccc  gtagaaaaga  tcaaaggatc   5460

ttcttgagat  ccttttttc  tgcgcgtaat  ctgctgcttg  caaacaaaaa  aaccaccgct   5520

accagcggtg  gtttgtttgc  cggatcaaga  gctaccaact  ctttttccga  aggtaactgg   5580

cttcagcaga  gcgcagatac  caaatactgt  ccttctagtg  tagccgtagt  taggccacca   5640

cttcaagaac  tctgtagcac  cgcctacata  cctcgctctg  ctaatcctgt  taccagtggc   5700

tgctgccagt  ggcgataagt  cgtgtcttac  cgggttggac  tcaagacgat  agttaccgga   5760

taaggcgcag  cggtcgggct  gaacggggggg  ttcgtgcaca  cagcccagct  tggagcgaac   5820

gacctacacc  gaactgagat  acctacagcg  tgagctatga  gaaagcgcca  cgcttcccga   5880

agggagaaag  gcggacaggt  atccggtaag  cggcagggtc  ggaacaggag  agcgcacgag   5940

ggagcttcca  ggggggaaacg  cctggtatct  ttatagtcct  gtcgggtttc  gccacctctg   6000

acttgagcgt  cgatttttgt  gatgctcgtc  aggggggcgg  agcctatgga  aaaacgccag   6060

caacgcggcc  ttttttacggt  tcctgggctt  ttgctggcct  tttgctcaca  tgttctttcc   6120

tgcgttatcc  cctgattctg  tggataaccg  tattaccgcc  tttgagtgag  ctgataccgc   6180

tcgccgcagc  cgaacgaccg  agcgcagcga  gtcagtgagc  gaggaagcgg  aag          6233
```

```
<210>  11
<211>  234
<212>  DNA
<213>  artificial

<220>
<223>  attP containing polynucleotide

<400>  11
gactagtact  gacggacaca  ccgaagcccc  ggcggcaacc  ctcagcggat  gccccggggc    60

ttcacgtttt  cccaggtcag  aagcggtttt  cgggagtagt  gccccaactg  gggtaacctt   120

tgagttctct  cagttggggg  cgtagggtcg  ccgacatgac  acaagggggtt  gtgaccgggg   180

tggacacgta  cgcgggtgct  tacgaccgtc  agtcgcgcga  gcgcgactag  taca         234


<210>  12
<211>  26
```

Integrase.ST25

```
<212>   DNA
<213>   artificial

<220>
<223>   Primer attB-for

<400>   12
taccgtcgac gatgtaggtc acggtc                                                 26
```